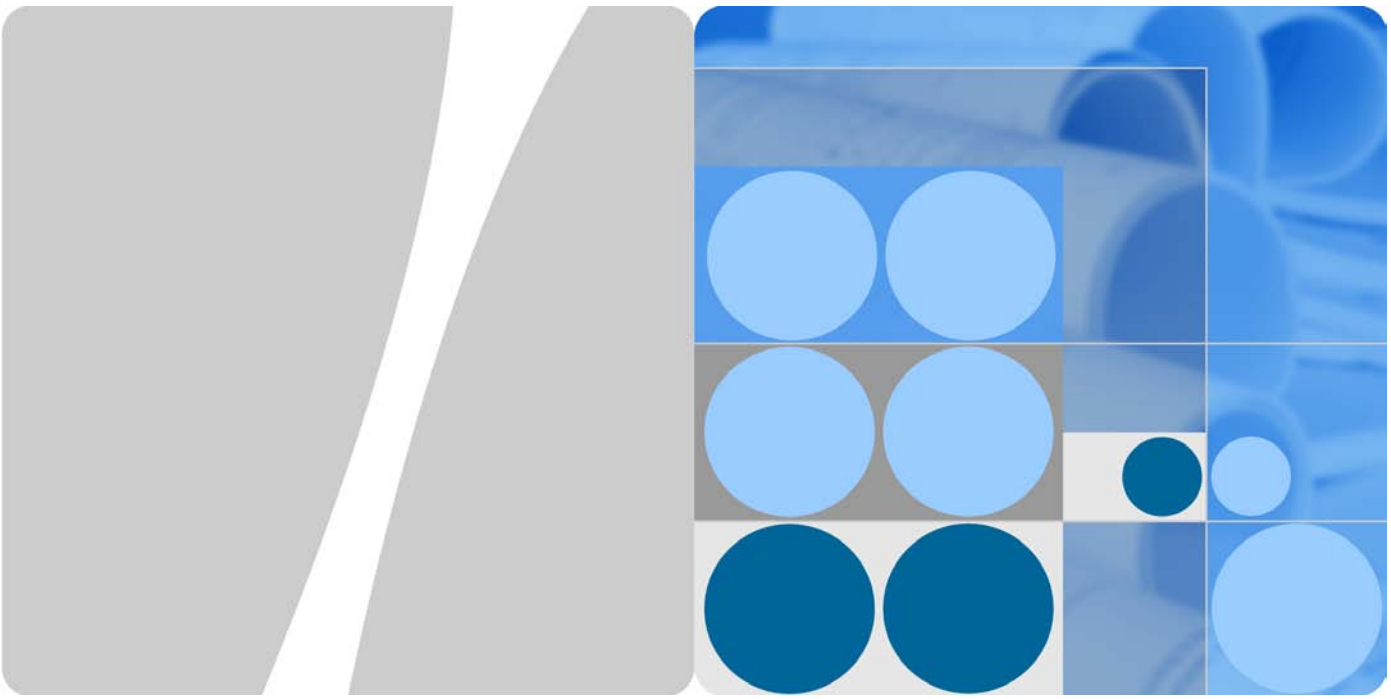


SmartAX MA5600T/MA5603T/MA5608T Multi-service Access Module V800R016C10

Feature Guide

Issue 02

Date 2015-12-30

HUAWEI TECHNOLOGIES CO., LTD.

Trademarks and Permissions


and other Huawei trademarks are trademarks of Huawei Technologies Co., Ltd.

All other trademarks and trade names mentioned in this document are the property of their respective holders.

Notice

The purchased products, services and features are stipulated by the contract made between Huawei and the customer. All or part of the products, services and features described in this document may not be within the purchase scope or the usage scope. Unless otherwise specified in the contract, all statements, information, and recommendations in this document are provided "AS IS" without warranties, guarantees or representations of any kind, either express or implied.

The information in this document is subject to change without notice. Every effort has been made in the preparation of this document to ensure accuracy of the contents, but all statements, information, and recommendations in this document do not constitute a warranty of any kind, express or implied.

Huawei Technologies Co., Ltd.

Address: Huawei Industrial Base
Bantian, Longgang
Shenzhen 518129
People's Republic of China

Website: http://www.huawei.com

Email: support@huawei.com

About This Document

Intended Audience

This document describes the key features (including ADSL,VDSL2, SHDSL, GPON, VoIP, ISDN, FoIP, MoIP, P2P Access, Layer 2 Protocol Handling, Layer 3 Features, VLAN, ACL, QoS, Multicast and security features) of the SmartAX MA5600T/MA5603T/MA5608T (hereinafter referred to as the MA5600T/MA5603T/MA5608T) in detail from the following aspects:

- Definition
- Purpose
- Specification
- Availability
- Principle
- Reference

After reading this document, you can learn about the definitions and purposes of the various features of the MA5600T/MA5603T/MA5608T, and also the support of these features by the MA5600T/MA5603T/MA5608T and the references on these features. In this way, you can know the feature list of the MA5600T/MA5603T/MA5608T and understand the implementation of these features on the MA5600T/MA5603T/MA5608T.

This document is intended for:

- Network planning engineers
- System maintenance engineers
- Configuration engineers
- NM administrators

Symbol Conventions

The following symbols may be found in this document. They are defined as follows

Symbol	Description
	Indicates an imminently hazardous situation which, if not avoided, will result in death or serious injury.
	Indicates a potentially hazardous situation which, if not avoided, could result in death or serious injury.

Symbol	Description
CAUTION	Indicates a potentially hazardous situation which, if not avoided, may result in minor or moderate injury.
NOTICE	Indicates a potentially hazardous situation which, if not avoided, could result in equipment damage, data loss, performance deterioration, or unanticipated results. NOTICE is used to address practices not related to personal injury.
NOTE	Calls attention to important information, best practices and tips. NOTE is used to address information not related to personal injury, equipment damage, and environment deterioration.

Update History

Overall updates between document issues are cumulative. Updates history of this document due to the updates of the product software, see the Feature Updates section of each chapter. The latest document issue contains all updates made in previous issues.

NOTE
V800R006C02 is the first version to record the update history of each feature. The feature updates between V800R006C02 and other earlier version are not included in this document. If you need the feature updates between V800R006C02 and other earlier version, please contact Huawei local support.

Issue 02 (2015-12-30)

Compared with issue 01 (2015-09-30) of V800R016C10, this issue has the following changes:

Position	Description
23.13.5 Ringing and CLIP Services for IP Z Interface Extension Feature	Added: Ringing and CLIP Services for IP Z Interface Extension Feature.
23.28 Configuring the IP Z Interface Extension Service	Modified: Configuring the IP Z Interface Extension Service.
1 Feature Specifications and Limitations	Modified: Feature Specifications and Limitations
Dual-Homing GPON Type B Protection Principles 2.13.6 Configuring GPON Type B Dual-Homing Protection	Added the Type B dual homing protection principle and configuration guide.
13.4 Service Flow	Added the Automatic Service Flow Creation principle and

Position	Description
	configuration guide.

Issue 01 (2015-09-30)

Compared with issue 02 (2015-07-28) of V800R016C00, this issue has the following changes:

Position	Description
23.13 IP Z Interface Extension	Added the IP Z interface extension.
23.15.8 Signaling Tracing	Added the voice service maintenance and diagnosis method of signaling tracing.
21.4 Configuring NAC-based Remote Software Commissioning Using GE Upstream Transmission	Added the configuring method through RN commands.
2.7.6 Energy Conservation	Added the Energy conservation.

Contents

1 Feature Specifications and Limitations

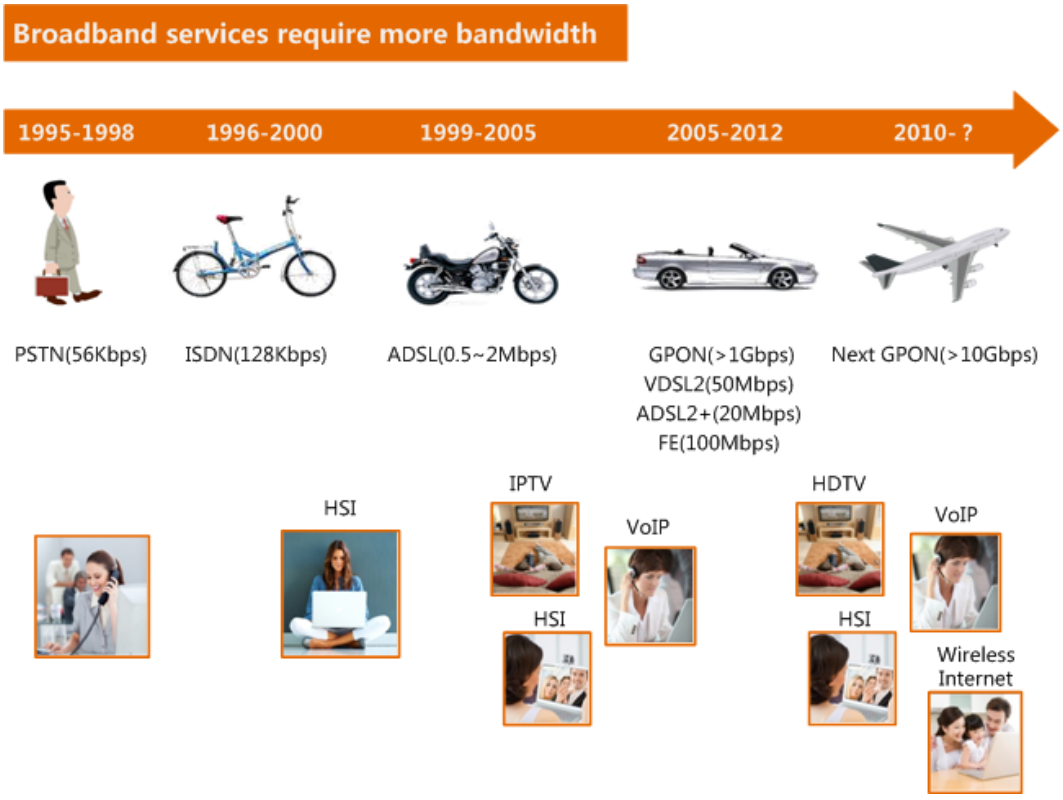
For detailed feature specifications and limitations, see *Feature Specifications and Limitations*.

2 GPON

About This Chapter

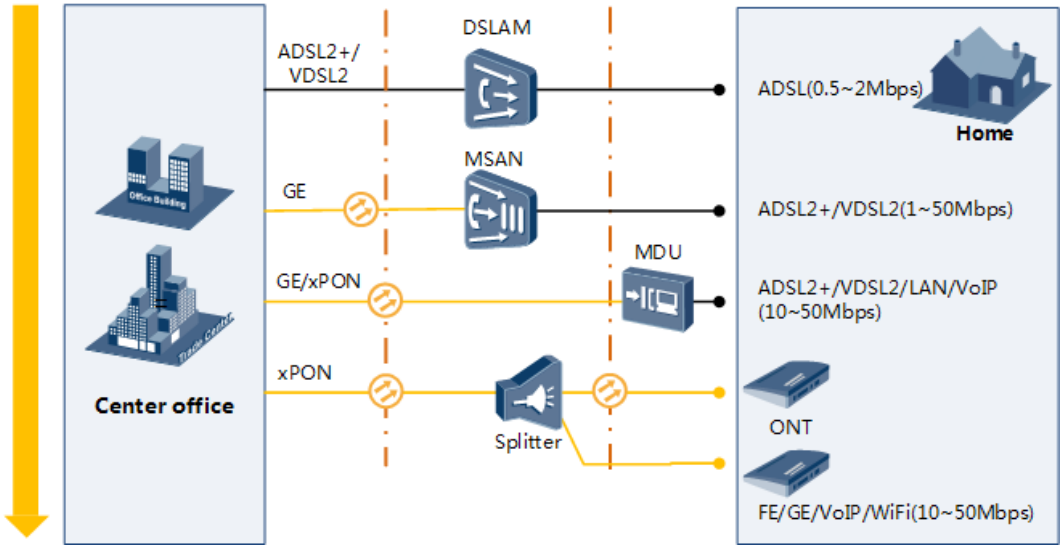
Gigabit passive optical network (GPON) is a PON technology that is standardized by the ITU-T Recommendations G.984.x. A GPON device supports high-bandwidth transmission. GPON effectively solves the bandwidth bottleneck problem in the twisted-pair access and meets users demands on high-bandwidth services.

2.1 Why Is GPON Required

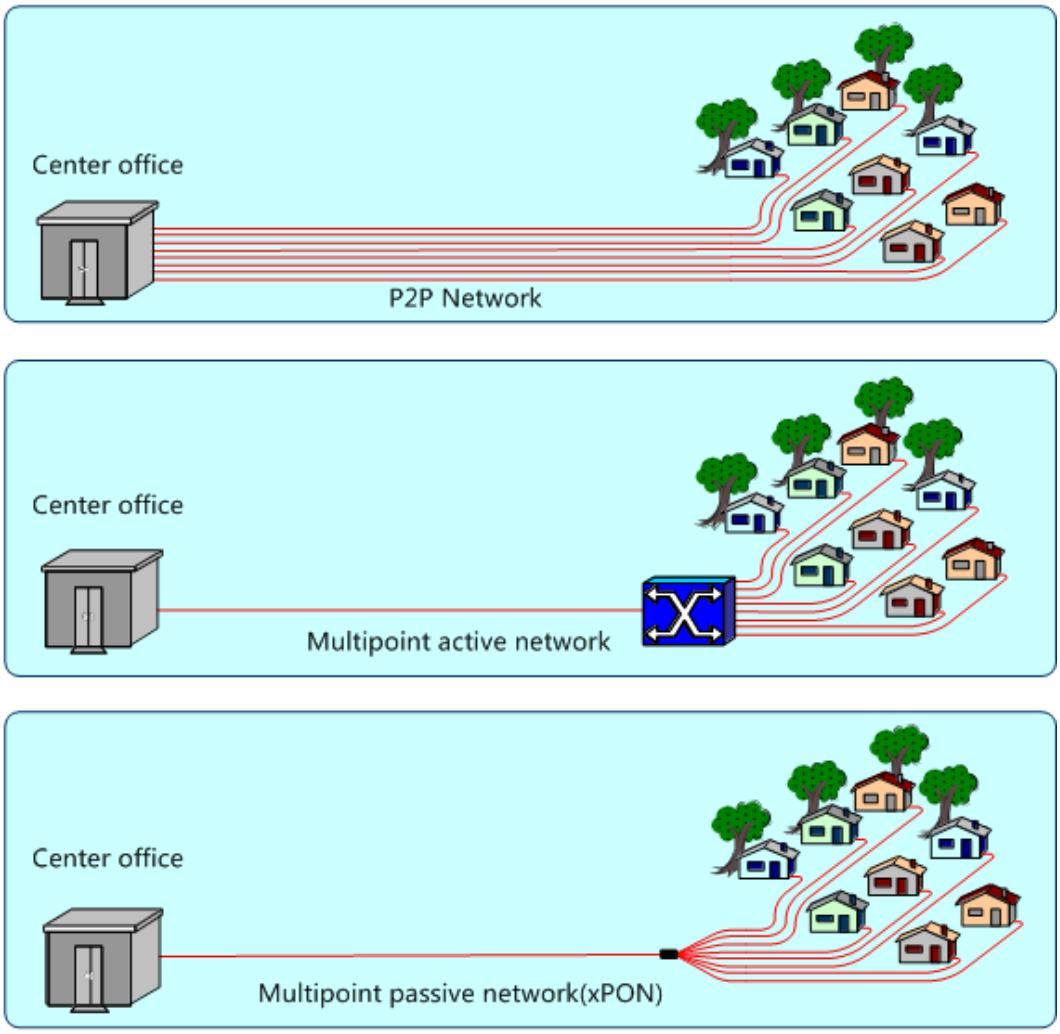


Access network evolution

The Application of optical fibers resolving transmission distance and bandwidth issues in twisted pair transmission.



GPON technology development



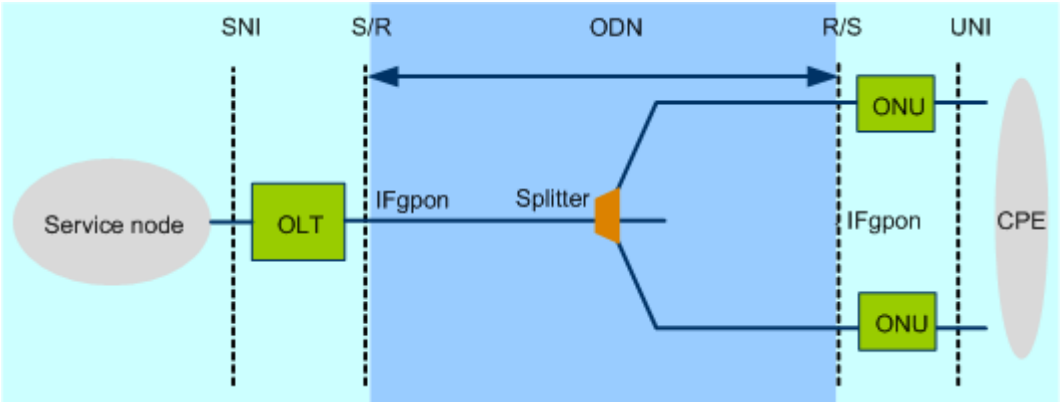
- Low-cost network
  - Savings of at least half of the fiber optic backbone
  - Compared to P2P, saving nearly half of the optical module
- Low maintenance costs
  - Passive Network save maintenance costs
  - Feature reduces the power consumption of passive cluster node
  - OAM management system reduces management difficulties
- Better quality of service
  - Features native support for multicast
  - Access high-bandwidth, and can evolution to more high bandwidth.

2.2 Introduction to GPON

What Is GPON

PON is a point to multi-point (P2MP) passive optical network. GPON, a type of PON technology, is defined by ITU-T Recommendation G.984.x. Figure 2-1 shows a GPON network.

Figure 2-1 GPON network



IFgpon: GPON interface

SNI: service node interface

UNI: user to network interface

CPE: customer premises equipment

- The optical line terminal (OLT) is an aggregation device located at the central office (CO) for terminating the PON protocol.
- Optical network units (ONUs)/Optical network terminal (ONTs) are located on the user side, providing various ports for connecting to user terminals. The OLT and ONUs are connected using an optical distribution network (ODN) for communication.
- The ODN is composed of passive optical components (POS), such as optical fibers, and one or more passive optical splitters. The ODN provides optical channels between the OLT and ONUs. It interconnects the OLT and ONUs and is highly reliable.

NOTE

The ODN network is passive, indicating that no optical amplifier or regenerator is deployed on the ODN network, thereby reducing maintenance costs of outdoor devices.

Why Is GPON Required

As the wide use of broadband services and fiber-in and copper-out development, carriers require a longer transmission reach, higher bandwidth, reliability, and lower operating expense (OPEX) on services. GPON supports the following functions to meet these requirements:

- Longer transmission distance: The transmission media of optical fibers covers up to 60 km coverage radius on the access layer, resolving transmission distance and bandwidth issues in twisted pair transmission.

- Higher bandwidth: Each GPON port can support a maximum transmission rate of 2.5 Gbit/s in the downstream direction and 1.25 Gbit/s in the upstream direction, meeting the usage requirements of high-bandwidth services, such as high definition television (HDTV) and outside broadcast (OB).
- Better user experience on full services: Flexible QoS measures support traffic control based on users and user services, implementing differentiated service provisioning for different users.
- Higher resource usage with lower costs: GPON supports a split ratio up to 1:128. A feeder fiber from the CO equipment room can be split to up to 128 drop fibers. This economizes on fiber resources and O&M costs.

2.3 Basic Concepts

GEM Frame

In the gigabit-capable passive optical network (GPON) system, a GPON encapsulation mode (GEM) frame is the smallest service-carrying unit and the basic encapsulation structure. All service streams are encapsulated into the GEM frame and transmitted over GPON lines. The service streams are identified by GEM ports and each GEM port is identified by a unique port ID. The port ID is globally allocated by the OLT. Therefore, the ONUs connected to the same OLT cannot use GEM ports that have the same port ID. A GEM port is used to identify the virtual service channel that carries the service stream between the OLT and the ONU. It is similar to the virtual path identifier (VPI)/virtual channel identifier (VCI) of the asynchronous transfer mode (ATM) virtual connection.

Figure 2-2 shows the GEM frame structure.

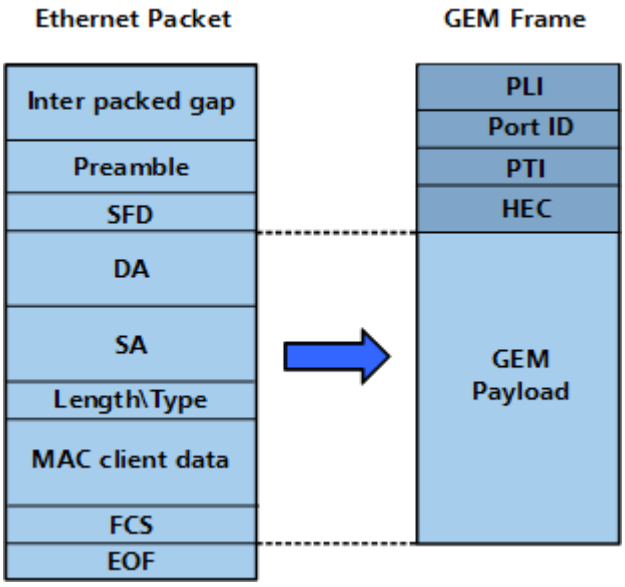
Figure 2-2 GEM frame structure

A GEM header consists of PLI, Port ID, PTI, and header error check (HEC) and is used for differentiating data of different GEM ports.

- PLI: indicates the length of data payload.
- Port ID: uniquely identifies a GEM port.
- PTI: indicates the payload type. It is used for identifying the status and type of data that is being transmitted, for example, whether the operation, administration and maintenance (OAM) message is being transmitted and whether data transmission is complete.
- HEC: ensures the forward error correction (FEC) function and transmission quality.
- Fragment payload: indicates the frame fragment.

The following section describes the GEM frame structure based on the mapping of the Ethernet service in GPON mode, as shown in Figure 2-3.

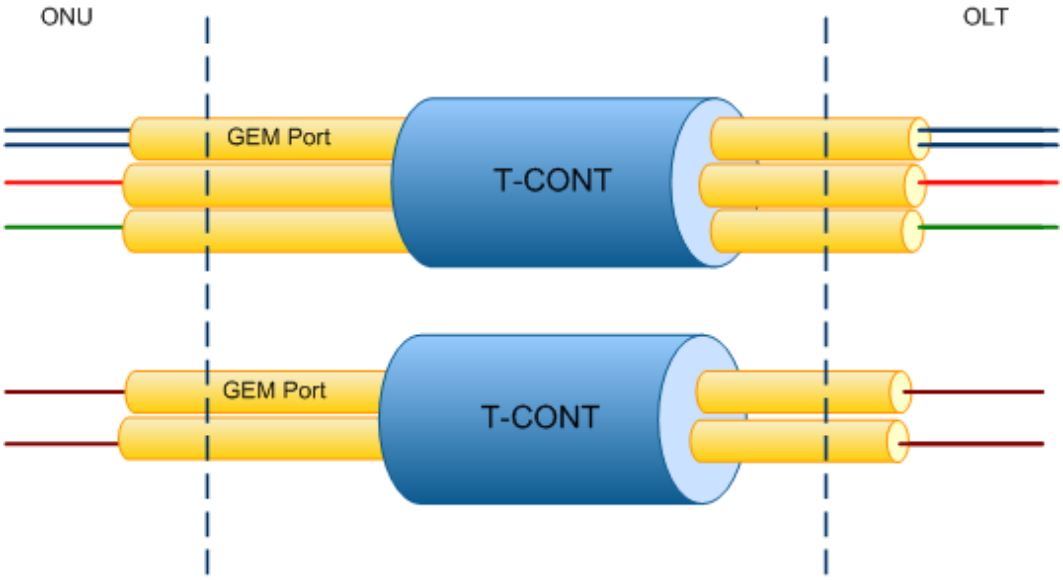
Figure 2-3 GEM frame structure



- The GPON system parses Ethernet frames and maps data into GEM payloads for transmission.
- Header information is automatically encapsulated into GEM frames.
- The mapping format is clear and has good compatibility.

T-CONT

Transmission container (T-CONT) is a service carrier in the upstream direction in the GPON system. All GEM ports are mapped to T-CONTs. Then service streams are transmitted upstream by means of OLT's dynamic bandwidth allocation (DBA) scheduling. T-CONT is the basic control unit of the upstream service stream in the GPON system. Each T-CONT is identified by Alloc-ID. The Alloc-ID is allocated by the GPON port of the OLT, and the T-CONTs used by ONUs connected to the same GPON port of OLT cannot have the same Alloc-IDs.kangyu



There are five types of T-CONT. T-CONT selection varies during the scheduling of different types of upstream service streams. Each T-CONT bandwidth type has its own quality of service (QoS) feature. QoS is mainly represented by the bandwidth guarantee, which can be classified into fixed, assured, non-assured, best-effort, and hybrid modes (corresponding to type 1 to type 5 listed in Table 2-1).

Table 2-1 T-CONT types

Bandwidth Type	T-CONT Type				
	Type 1	Type 2	Type 3	Type 4	Type 5
Fixed Bandwidth	X	No	No	No	X
Assured Bandwidth	No	Y	Y	No	Y
Maximum Bandwidth	Z = X	Z = Y	Z > Y	Z	Z ≥ X + Y
Description	• The fixed bandwidth is reserved for specific ONUs or specific services on ONUs. It cannot be used by other ONUs	• The assured bandwidth is available at any time required by an ONU. When the bandwidth required	• This type is the combination of the assured bandwidth and maximum bandwidth. The system assures some bandwidt	• This type is the maximum bandwidth that can be used by an ONU, fully providing the bandwidth required	This type is the combination of the fixed, assured, and maximum bandwidth. It supports the following functions: • Reserves bandwidth for subscribers and the

Bandwidth Type	T-CONT Type				
	Type 1	Type 2	Type 3	Type 4	Type 5
	even if no upstream service streams are carried on the specific ONUs. • It applies to services that are sensitive to service quality. The services can be TDM or VoIP services.	by the service streams on the ONU is smaller than the assured bandwidth, the system can use the DBA mechanism to allocate the remaining bandwidth to services on other ONUs. • Because DBA is required, this type provides a less real-time performance compared with the fixed bandwidth.	h for subscribers and allows subscribers to preempt bandwidth. However, the total used bandwidth cannot exceed the maximum configured bandwidth. • It applies to VoIP services.	by the ONU. • It applies to IPTV and other high-speed Internet services.	bandwidth cannot be preempted by other subscribers. • Provides the bandwidth to an ONU at any time when required • Allow subscribers to preempt some bandwidth. (The total used bandwidth cannot exceed the maximum configured bandwidth.)

NOTE

In Table 2-1, X indicates the fixed bandwidth value, Y indicates the assured bandwidth value, Z indicates the maximum bandwidth value, and No indicates not involved.

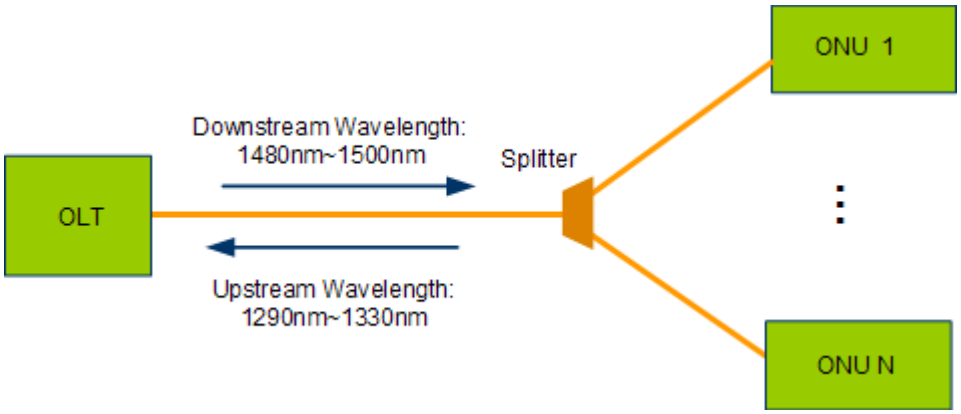
2.4 GPON System Overview

Introduction to the GPON System

Mainstream PON technologies include broadband passive optical network (BPON), Ethernet passive optical network (EPON), and gigabit passive optical network (GPON). Adopting the ATM encapsulation mode, BPON is mainly used for carrying ATM services. With the obsolescence of the ATM technology, BPON also drops out. EPON is an Ethernet passive optical network technology. GPON is a gigabit passive optical network technology and is to date the most widely used mainstream optical access technology.

Figure 2-4 shows the working principle of the GPON network.

Figure 2-4 Working principle of the GPON network

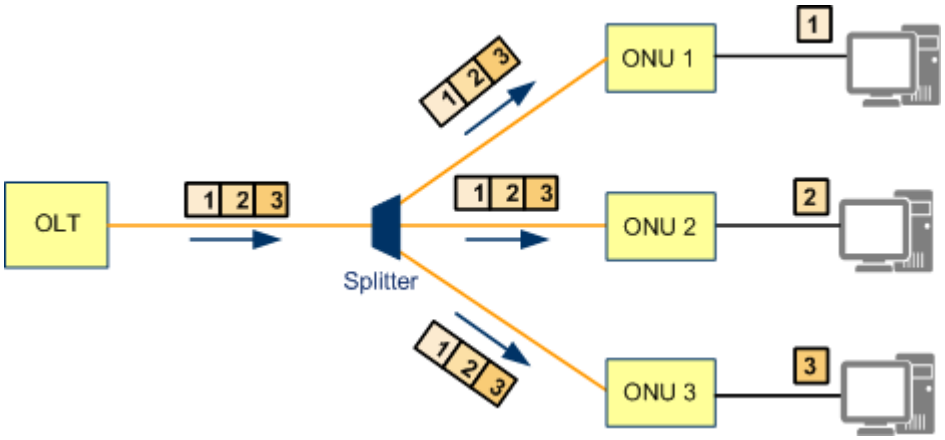


- In the GPON network, the OLT is connected to the optical splitter through a single optical fiber, and the optical splitter is then connected to ONUs. Different wavelengths are adopted in the upstream and downstream directions for transmitting data. Specifically, wavelengths range from 1290 nm to 1330 nm in the upstream direction and from 1480 nm to 1500 nm in the downstream direction.
- The GPON adopts WDM to transmit data of different upstream/downstream wavelengths over the same ODN. Data is broadcast in the downstream direction and transmitted in the TDMA mode (based on timeslots) in the upstream direction.

GPON Downstream Transmission

All data is broadcast to all ONUs from the OLT. The ONUs then select and receive their respective data and discard the other data. Figure 2-5 shows the details.

Figure 2-5 Downstream communication principle of GPON



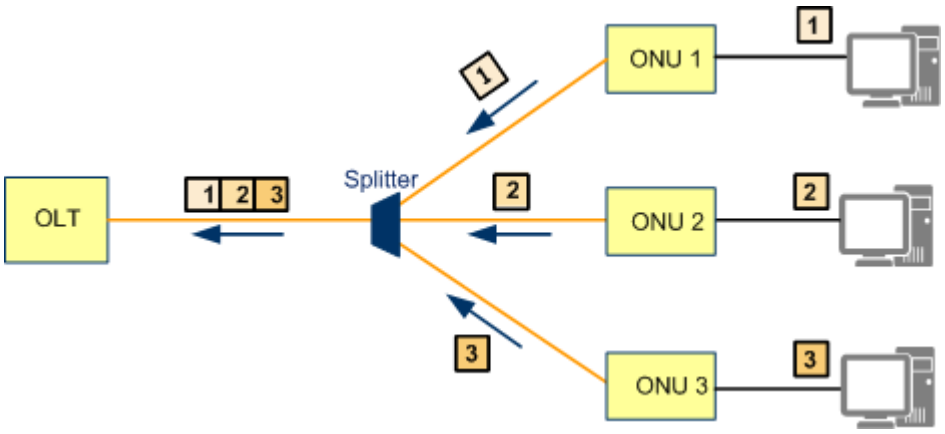
Main features:

- Supports point-to-multipoint (P2MP) multicast transmission.
- Broadcasts the same data to all ONUs and differentiates ONU data by GEM port ID.
- Allows an ONU to receive the desired data by ONU ID.

GPON Upstream Transmission

In the upstream direction, each ONU can send data to the OLT only in the timeslot permitted and allocated by the OLT. This ensures that each ONU sends data in a given sequence, avoiding upstream data conflicts. Figure 2-6 shows the details.

Figure 2-6 Upstream communication principle of GPON



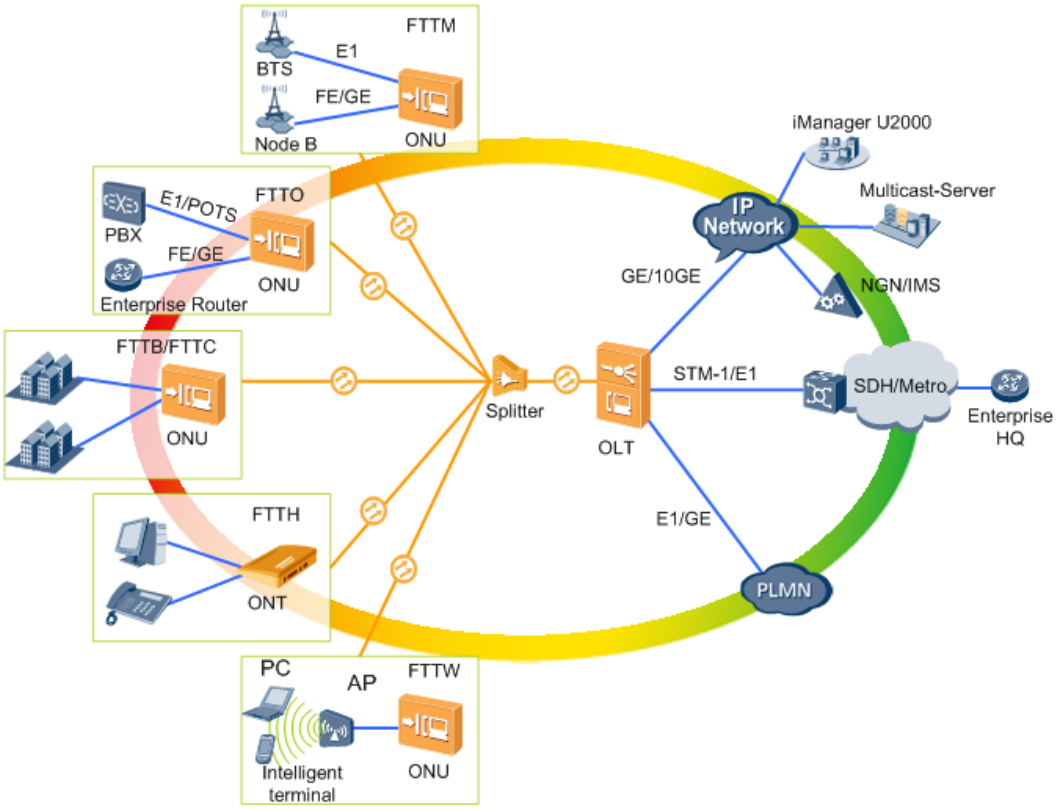
Main features:

- Supports time division multiple access (TDMA).
- Transits data on an exclusive timeslot.
- Couples optical signals on an optical splitter.
- Detects and prevents collisions through ranging.

2.5 GPON Networking Applications

GPON is a passive optical transmission technology that applies in FTTx solutions, including fiber to the building (FTTB), fiber to the curb (FTTC), fiber to the door (FTTD), fiber to the home (FTTH), fiber to the mobile base station (FTTM), fiber to the office (FTTO), and fiber to the WLAN (FTTW), for voice, data, video, private line access, and base station access services. Figure 2-7 shows FTTx networking applications.

Figure 2-7 FTTx networking applications



The FTTx network applications in GPON access have the following in common: The data, voice, and video signals of terminal users are sent to ONUs, where the signals are converted into Ethernet packets and then transmitted over optical fibers to the OLT using the GPON uplink ports on the ONUs. Then, the Ethernet packets are forwarded to the upper-layer IP network using the uplink port on the OLT.

- FTTB/FTTC: The OLT is connected to ONUs in corridors (FTTB) or by the curb (FTTC) using an optical distribution network (ODN). The ONUs are then connected to user terminals using xDSL. FTTB/FTTC is applicable to densely-populated residential communities or office buildings. In this scenario, FTTB/FTTC provides services of certain bandwidth for common users.
- FTTD: uses existing access media at user homes to resolve drop fiber issues in FTTH scenarios.
- FTTH: The OLT connects to ONTs at user homes using an ODN network. FTTH is applicable to new apartments or villas in loose distribution. In this scenario, FTTH provides services of higher bandwidth for high-end users.

- FTTM: The OLT is connected to ONUs using an ODN network. The ONUs are then connected to wireless base stations using E1. The OLT connects wireless base stations to the core IP bearer network using optical access technologies. This implementation mode is not only simpler than traditional SDH/ATM private line technologies, but also drives down the costs of base station backhaul. FTTM is applicable to reconstruction and capacity expansion of mobile bearer networks. In this scenario, FTTM converges the fixed network and the mobile network on the bearer plane.
- FTTO: The OLT is connected to enterprise ONUs using an ODN network. The ONUs are connected to user terminals using FE, POTS, or Wi-Fi. QinQ VLAN encapsulation is implemented on the ONUs and the OLT. In this way, transparent and secure data channels can be set up between the enterprise private networks located at different places, and therefore the service data and BPDUs between the enterprise private networks can be transparently transmitted over the public network. FTTO is applicable to enterprise networks. In this scenario, FTTO implements TDM PBX, IP PBX, and private line service in the enterprise intranets.
- FTTW: The OLT connects to ONUs using an ODN network, the ONUs connect to access points (APs) using GE for WLAN traffic backhaul. FTTW is the trend in Wi-Fi construction.

2.6 GPON Principles

2.6.1 GPON Service Multiplexing

GPON encapsulation mode (GEM) ports and transmission containers (T-CONTs) divide a PON network into virtual connections for service multiplexing.

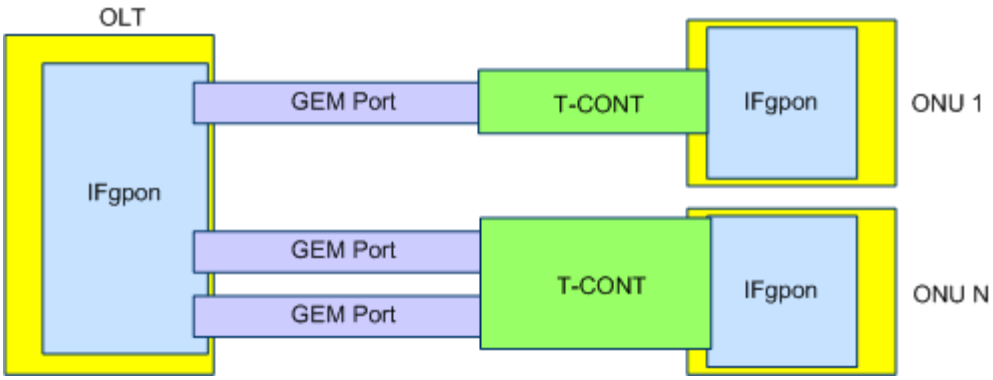
- Each GEM port can carry one or more types of service stream. After carrying service streams, a GEM port must be mapped to a T-CONT before upstream service scheduling. Each ONU supports multiple T-CONTs that can have different service types.
- A T-CONT can be bound to one or more GEM ports, depending on customers' data plan. On the OLT, GEM ports are demodulated from the T-CONT and then service streams are demodulated from the GEM port payload for further processing.

Service Mapping Relationships

- In the upstream direction,
  - An ONU sends Ethernet frames to GEM ports based on configured mapping rules between service ports and GEM ports. Then, the GEM ports encapsulate the Ethernet frames into GEM packet data units (PDUs) and add these PDUs to T-CONT queues based on mapping rules between GEM ports and T-CONT queues. Then, the T-CONT queues use timeslots for upstream transmission to send GEM PDUs to the OLT.
  - The OLT receives the GEM PDUs and obtains Ethernet frames from them. Then, the OLT sends Ethernet frames from a specified uplink port based on mapping rules between service ports and uplink ports.

  Figure 2-8 shows GPON service mapping relationships in the upstream direction.

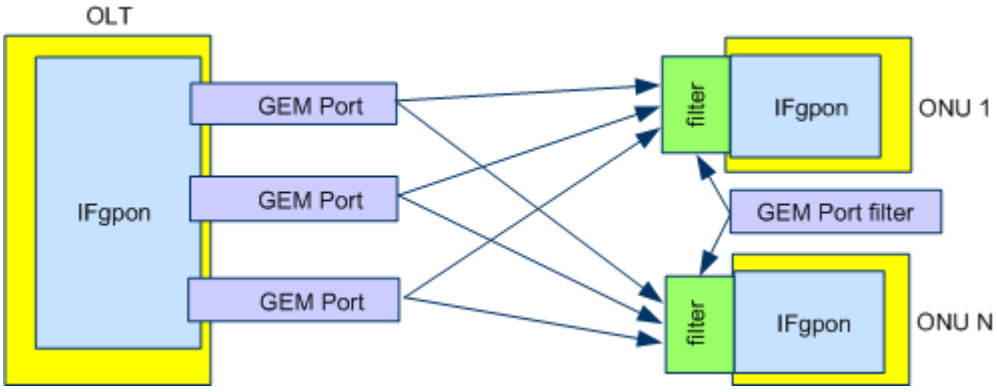
Figure 2-8 GPON service mapping relationships in the upstream direction



- In the downstream direction,
  - The OLT sends Ethernet frames to the GPON service processing module based on configured mapping rules between service ports and uplink ports. The GPON service processing module then encapsulates the Ethernet frames into GEM PDUs for downstream transmission using a GPON port.
  - GPON transmission convergence (GTC) frames containing GEM PDUs are broadcast to all ONUs connected to the GPON port.
  - The ONU filters the received data according to the GEM port ID contained in the GEM PDU header and retains the data only belonging to the GEM ports of this ONU. Then, the ONU decapsulates the data to Ethernet frames and sends them to end users using service ports.

  Figure 2-9 shows GPON service mapping relationships in the downstream direction.

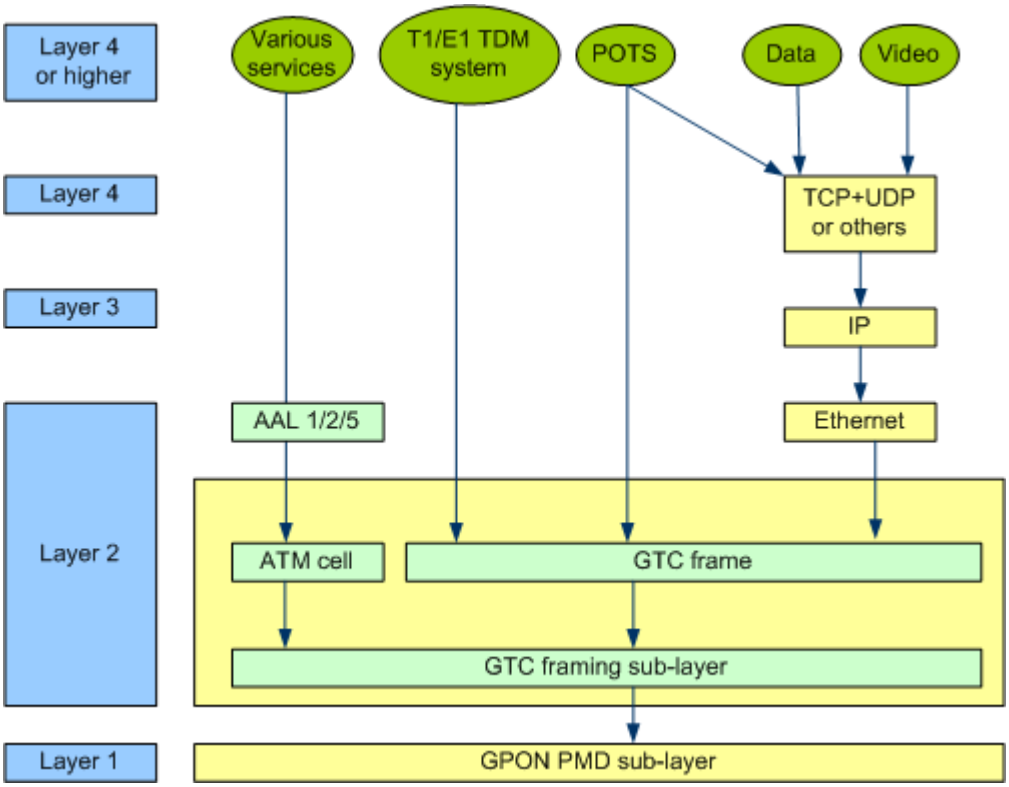
Figure 2-9 GPON service mapping relationships in the downstream direction



2.6.2 GPON Protocol Stacks

ITU-T Recommendation G.984.3 defines a new set of frame structures, which consider traditional voice, video, and Ethernet packets as payloads of GPON frames. Figure 2-10 shows the structure of GPON protocol stacks.

Figure 2-10 Structure of GPON protocol stacks



GPON protocol stacks involve the physical medium dependent (PMD) layer and GPON transmission convergence (GTC) layer.

PMD Layer

The GPON PMD layer corresponds to the GPON interfaces between OLTs and ONUs. Parameter values of the GPON interfaces specify the maximum reach and split ratio for a GPON system.

GTC Layer

The GTA layer is used to encapsulate payloads using ATM cells or GEM frames, and GEM frames are commonly used in GPON systems. GEM frames can carry Ethernet, POTS, E1, and T1 cells.

GTC is the core GPON layer, where media access is controlled for upstream service flows and ONUs are registered. Ethernet frame payloads are encapsulated into GEM frames and then packetized as GTC frames. These GTC frames are converted to binary codes for transmission based on interface parameters configured at the physical layer. The process is reversal on the receive end. Specifically, the receive end decapsulates the data to obtain GTC frames, GEM frames, and then payloads for data transmission.

The GTC layer is classified as TC adaptation sub-layer and GTC framing sub-layer by structure.

- The TC adaptation sub-layer involves the ATM, GEM TC, and optical network terminal management and control interface (OMCI) adapters and dynamic bandwidth assignment (DBA) control module. ATM and GEM TC adapters identify OMCI channels by virtual path identifier (VPI)/virtual channel identifier (VCI) or GEM port ID. OMCI adapters

can exchange OMCI channel data with the ATM and GEM TC adapters and send the OMCI channel data to OMCI entities. The DBA control module is a common functional module, which generates ONU reports and controls DBA allocation.

- On the GTC framing sub-layer, GTC frames include GEM blocks, PLOAM blocks, and embedded OAM blocks. The GTC framing sub-layer supports the following functions:
  - Multiplexes and demultiplexes data. Specifically, the GTC framing sub-layer multiplexes PLOAM and GEM data into downstream TC frames based on the boundary information specified in the frame header. In addition, the GTC framing sub-layer demultiplexes PLOAM and GEM data from upstream TC frames based on frame header instructions.
  - Generates frame headers and decodes data. The GTC framing sub-layer generates the TC header of downstream frames in a specified format and decodes the frame header of upstream frames. In addition, the GTC framing sub-layer terminates the embedded OAM data encapsulated into the GTC header and uses the OAM data to control this sub-layer.
  - Routes data internally based on alloc-IDs. The GTC framing sub-layer routes the data sent by or to the GEM TC adapters based on internal alloc-IDs.

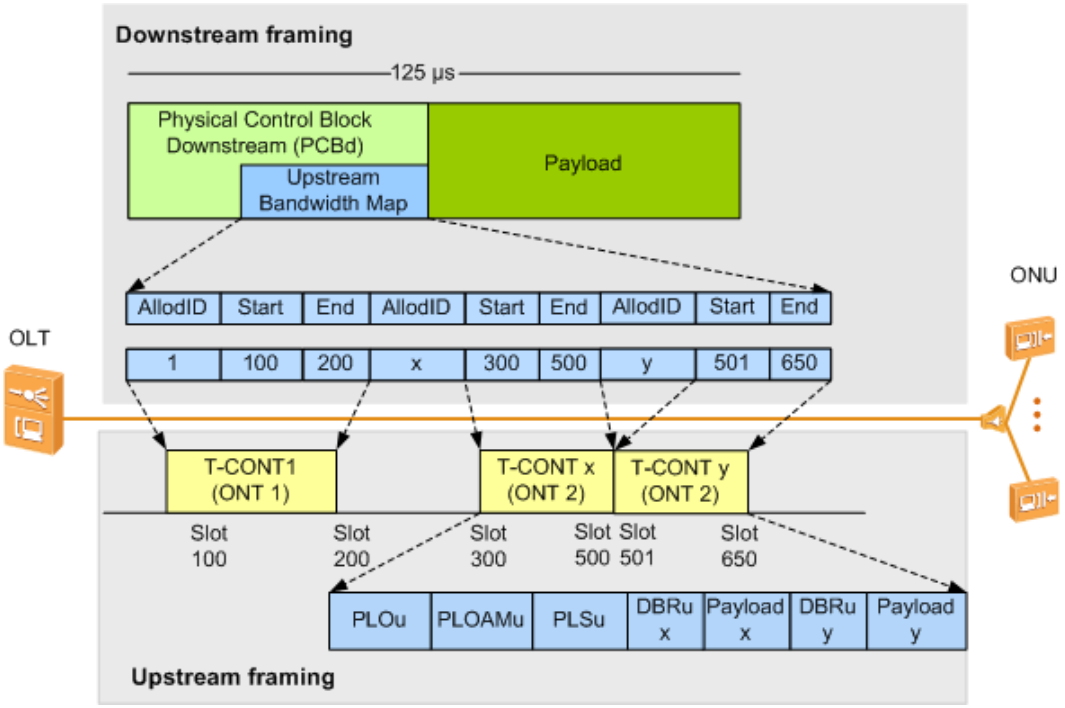
The GTC layer consists of plane C/M and plane U based on functions.

- The protocol stacks of plane C/M include embedded OAM, PLOAM, and OMCI. Embedded OAM and PLOAM channels are used for managing PMD and GTC sub-layer functions. OMCI provides a unified system for upper-layer sub-layer management.
  - Embedded OAM channels are defined in GTC frame headers for determining bandwidths, exchanging data, and dynamically allocating bandwidths.
  - Dedicated space is reserved in GTC frames for format-based PLOAM channels. The PLOAM channels carry the PMD and GTC management information that does not pass through the embedded OAM block.
  - OMCI channels are used for managing services.
- Service flows on plane U are identified based on service flow types (ATM or GEM) and port ID/VPI. Port IDs identify GEM service flows and VPIs identify ATM service flows. In T-CONTs, bandwidths are allocated and QoS is controlled using the timeslots that can be adjusted.

2.6.3 GPON Frame Structure

Figure 2-11 shows the GPON frame structure.

Figure 2-11 GPON frame structure



Upstream GPON Frame

An upstream GPON frame has a fixed length of 125 µs. Each upstream frame contains the content carried by one or more T-CONTs. All ONUs connected to a GPON port share the upstream bandwidth

- All ONUs connected to a GPON port send their data upstream at their own timeslots according to bandwidth map (BWmap) requirements.
- Each ONU reports the status of data to be sent to the OLT using upstream frames. Then, the OLT uses DBA to allocate upstream timeslots to ONUs and sends updates in each frame.

In Figure 2-11, an upstream GPON frame consists of the physical layer overhead upstream (PLOu), PLOAM upstream (PLOAMu), power level sequence upstream (PLSu), dynamic bandwidth report upstream (DBRu), and payload fields, as described in Table 2-2.

Table 2-2 Field description for an upstream GPON frame

Field	Description	Function
PLOu	Upstream physical layer overhead	Used for frame alignment, synchronization, and identification for an ONU.
PLOAMu	PLOAM messages of upstream data	Used for reporting ONU management messages, including maintenance and management status. This field may not be contained in a frame but must be negotiated.
PLSu	Upstream power level sequence	Used by ONUs for adjusting optical port power. This field may not be contained in a frame but must be

Field	Description	Function
		negotiated.
DBRu	Upstream dynamic bandwidth report	Used for reporting the T-CONT status to apply for bandwidth next time and for allocating dynamic bandwidths. This field may not be contained in a frame but must be negotiated.
Payload	Payload user data	Can be a DBA status report or data frame. If this field is a data frame, this field consists of a GEM header and frames.

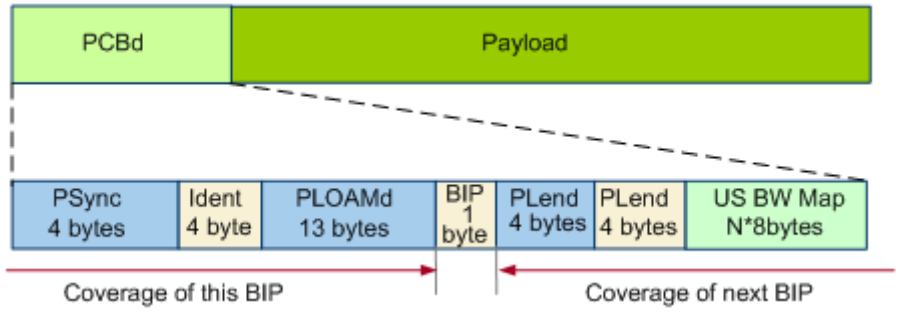
Downstream GPON Frame

A downstream GPON frame has a fixed length of 125 μs and comprises physical control block downstream (PCBd) and payload. PCBd mainly consists of the GTC header and BWmap. The OLT broadcasts PCBd to all ONUs. Then, the ONUs receive the PCBd and perform operations based on the information contained in PCBd.

- The GTC header is used for frame delimitation, synchronization, and forward error correction (FEC).
- The BWMap field notifies every ONU of upstream bandwidth allocation. It specifies the start and end upstream timeslots for the T-CONTs of each ONU, ensuring that all ONUs send data using the timeslots specified by the OLT to prevent data conflict.

Figure 2-12 shows the structure of the PCBd shown in Figure 2-11.

Figure 2-12 PCBd structure



PCBd contains PSync, Ident, PLOAMd, BIP, PLend, and US BW Map fields, where US BW Map is the upstream bandwidth mapping sent by the OLT for each T-CONT. Table 2-3 describes each field.

Table 2-3 PCBd field description

Field	Description	Function
PSync	Physical synchronization domain, frame synchronization information	Used by ONUs to specify the start of each frame.

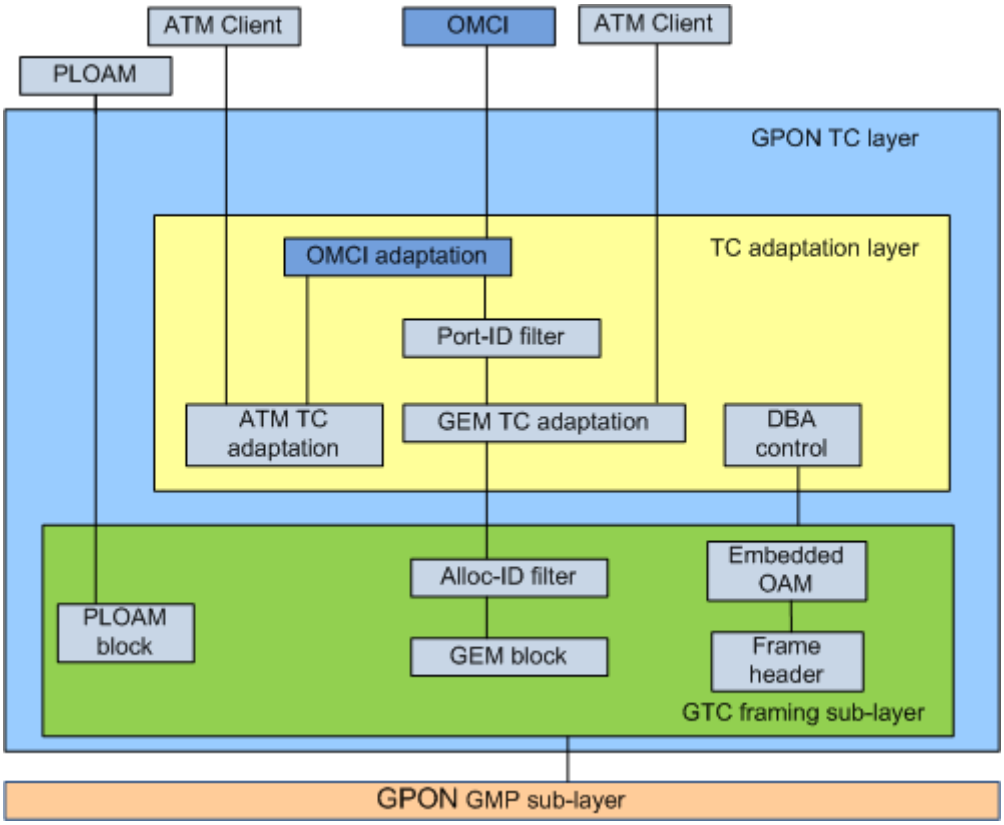
Field	Description	Function
Ident	Identification domain	Used for sorting a frame in the frames of the same type in length sequence.
Downstream PLOAM (PLOAMd)	PLOAM messages of downstream data	Used for reporting ONU management messages, including maintenance and management status. This field may not be contained in a frame but must be negotiated.
BIP	Bit interleaved parity	Used for performing a parity check for all bytes between two BIP fields (excluding the preamble and delimit) to monitor error codes.
PLend	Length of downstream payloads	Used for specifying the length of the BWmap field.
Upstream bandwidth map (US BW Map)	Upstream bandwidth mapping	Used by the OLT for sending the upstream bandwidth mapping to each T-CONT. The BWmap specifies the start and end times for each T-CONT in transmitting data.

2.6.4 OMCI

Basic Concepts

OMCI is a type of ITU-T Recommendation G.984.4-compliant configuration and transmission channel, which is used to transmit OMCI messages over dedicated ATM PVCs or GEM ports established between an OLT and an ONT. The OMCI messages are used for discovering ONTs for management and control.

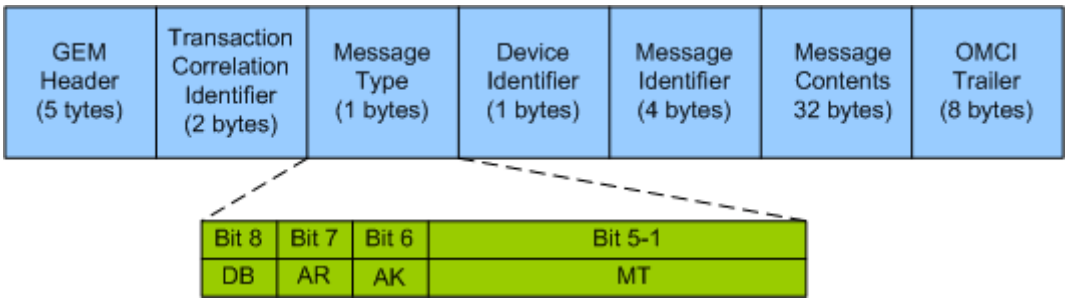
OMCI Position in GPON Protocol Stacks



OMCI Message Format

OMCI messages are strictly limited in length and format. Specifically, the length is consistently 53 bytes and the length of the OMCI data unit is 48 bytes. Figure 2-13 shows the OMCI message format.

Figure 2-13 OMCI message format



- GEM Header: includes GEM payload, GEM port ID, payload type indicator (PTI), and header error control (HEC).
- Transaction Correlation Identifier: The value of this field must be the same in a request and the response to this request. The highest order of this field indicates the priority of an OMCI message. Value 0 indicates a low priority and value 1 indicates a high priority.

- Message type:
  - DB: a destination bit, which is consistently 0.
  - AR: an acknowledge request, indicating whether an OMCI message requires the response from the peer end. Value 0 indicates that the response is not required and value 1 indicates that the response is required.
  - AK: acknowledgement, indicating whether an OMCI message is a response. Value 0 indicates not and value 1 indicates yes.
  - MT: message type, which supports up to 32 message types, including Create, Delete, Set, Get, and MIB upload. In ITU-T Recommendation G.984.4, message types 4 through 28 are used and other message types are reserved.
- Device identifier: The value of this field is consistently **0xA**.
- Message Identifier: a 2-byte entity or instance ID.
- Message Contents: packet payload.
- OMCI trailer: Two bytes are consistently 0, two bytes are packet length 0x28, and four bytes are CRCs.

OMCI Management

The OLT controls the ONT using the OMCI. The OMCI protocol allows the OLT to:

- Establish and release connections with the ONT.
- Manage the UNIs on the ONT.
- Request configuration information and performance statistics.
- Autonomously inform the system administrator of events, such as link failures.

The OMCI protocol runs over a GEM connection between the OLT controller and the ONT controller. The GEM connection is established during ONT initialization. The OMCI protocol is asynchronous: the OLT controller is the master and the ONT controller is the slave. A single OLT controller using multiple protocol instances over separate control channels can control multiple ONTs.

The OLT manages the ONT using OMCI in the following aspects:

- Configuration management: Controls and identifies the ONT, and collects data from and provides data to the ONT.
- Fault management: Supports limited fault management. Most of the operations are limited to failure indication.
- Performance management: Collects and queries performance statistics.
- Security management: Enables/Disables downstream encryption.

2.7 Key GPON Techniques

A series of key GPON techniques are applied to improve bandwidths and stabilities of GPON lines. This section describes key GPON techniques.

Key GPON techniques include:

- Ranging
- Burst optical or electrical technology
- DBA

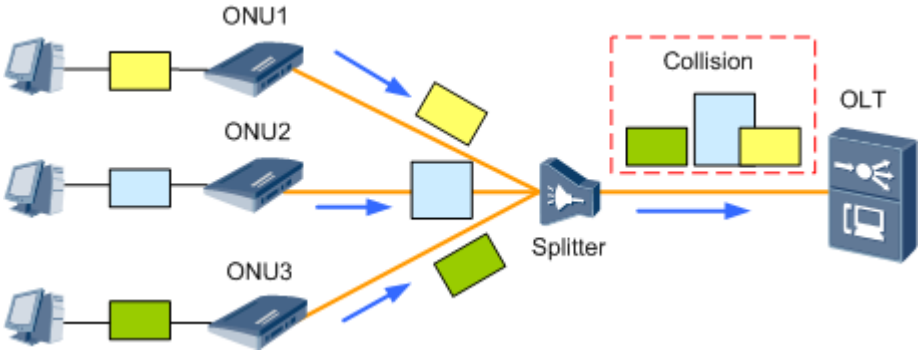
- FEC
- Line encryption

2.7.1 Ranging

Why Is Ranging Required

The logic reaches from ONUs to an OLT vary. Therefore, the time required for transmitting optical signals over optical fibers is different and the times when the ONUs receive optical signals is different. In addition, the round trip delays (RTDs) between an OLT and ONUs also vary depending on time and environment. Therefore, collisions may occur when ONU sends data in TDMA mode (in this mode, only one of the ONUs connecting to a PON port sends data at a moment), as shown in Figure 2-14. The OLT must precisely measure the distances between itself and each ONU to provide a proper timeslot for converged upstream data from all ONUs to prevent data conflict. In this way, the OLT controls the time for each ONU to send data upstream.

Figure 2-14 Cell transmission without ranging



Ranging Principles

Ranging process is as follows:

- The OLT starts ranging for an ONU when the ONU registers with the OLT for the first time and obtains the round trip delay (RTD) of the ONU. Based on the RTD, the OLT calculates the physical reach of this ONU.
- The OLT specifies a proper equalization delay (EqD) for the ONU based on the physical reach.

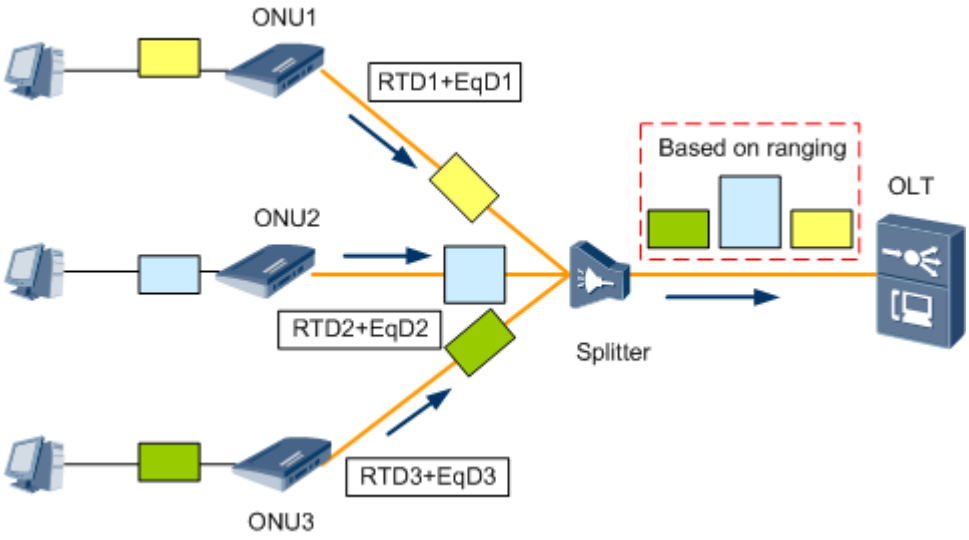
NOTE

The OLT requires a quiet zone during ranging to pause the upstream transmission channel of the ONUs connected to it. The quiet zone is implemented by emptying BWmap so that no timeslot is allocated for data transmission.

Ranging Results

RTD and EqD synchronize data frames sent by all ONUs, preventing data conflict on optical splitters. In this way, all ONUs locate at the same logic reach and they send data at specified timeslots, thereby preventing upstream cell conflict.

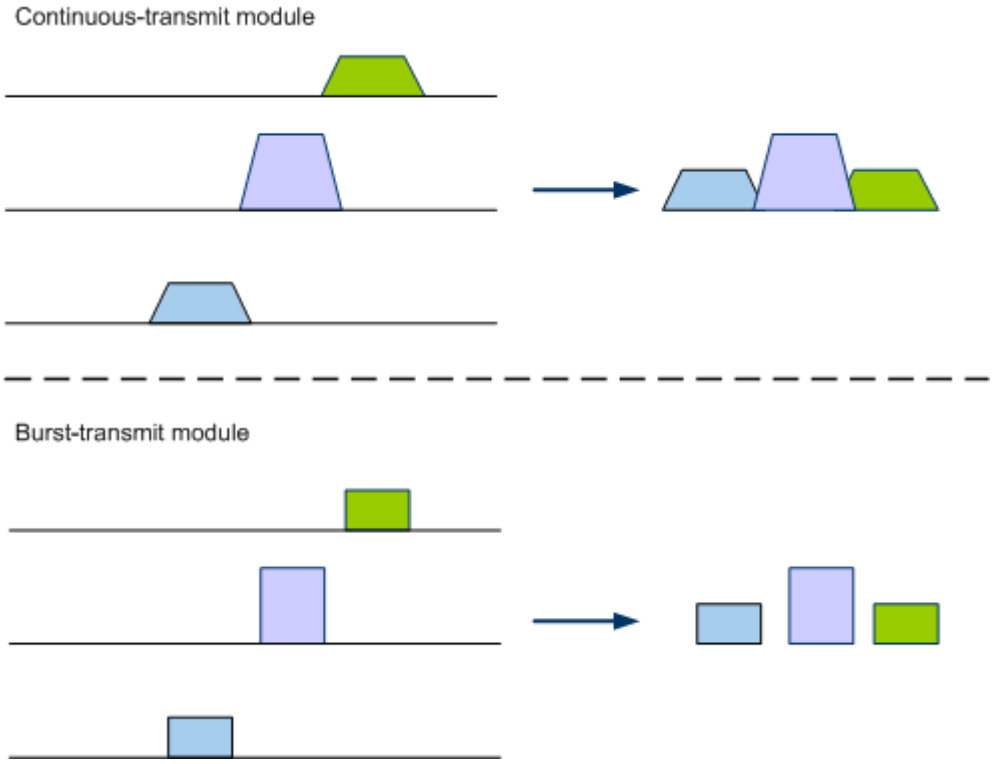
Figure 2-15 Cell transmission with ranging



2.7.2 Burst Optical/Electrical Technology

TDMA is used in GPON upstream direction. An ONU transmits data only within the allocated timeslots. In the timeslots that are not allocated to it, the ONU immediately disables the transmission of its optical transceiver to prevent other ONUs from being affected. The OLT then receives the upstream data from each ONU in a burst manner based on timeslots. Therefore, both OLT and ONU optical modules must support burst receive and transmit function to ensure normal running of the GPON system. Figure 2-16 shows the burst transmit function supported by ONU optical modules, and Figure 2-17 shows the burst receive function supported by OLT optical modules.

Figure 2-16 Burst transmit function supported by ONU optical modules

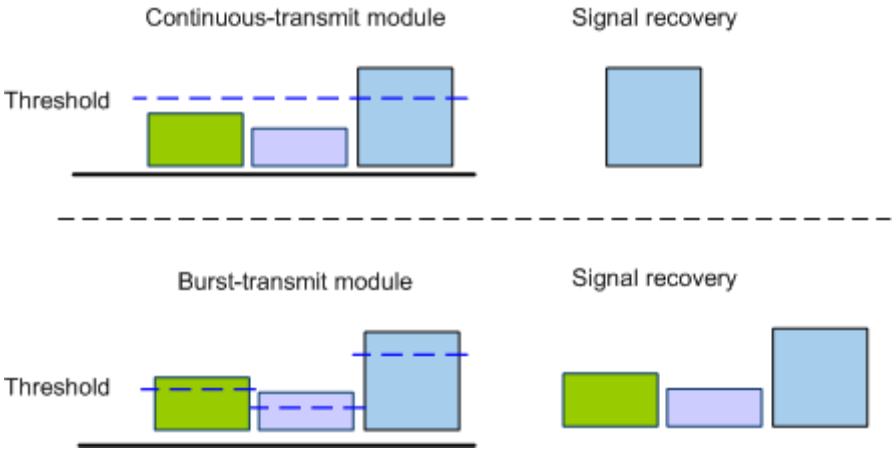


Ranging can be implemented to prevent cells transmitted by different ONUs from conflicting with each other on the OLT. However, the ranging accuracy is ± 1 bit and the cells transmitted by different ONUs have a protection time of several bits (not a multiple of 1 bit). If the ONU optical modules do not support the burst receive and transmit function, the transmitted signals overlap and distortion occurs.

NOTE

In the GPON system, all data is broadcast downstream to ONUs. The transmission requires OLT optical modules to transmit optical signals continuously and ONU optical modules to receive optical signals continuously. Therefore, these optical modules are not required to support the burst receive and transmit function.

Figure 2-17 Burst receive function supported by OLT optical modules



- The distance from each ONU to the OLT varies and therefore the optical signal attenuation varies for each ONU. As a result, the power and level of packets received by an OLT at different timeslots various.
- If the OLT optical modules do not support the burst receive and transmit function, an error occurs when the optical signals sent by the ONU with a long transmission distance and large optical attenuation are recovered on the OLT because the optical power level is less than the threshold (only the signals with the optical power level greater than the threshold can be recovered). Dynamic threshold adjustment enables the OLT to dynamically adjust the threshold for optical power levels based on the strengths of signals received by the OLT. This ensures that all ONU signals can be recovered.

2.7.3 DBA

In the GPON system, the OLT controls an ONU's upstream data traffic by sending authorization signals to the ONU. PON requires an effective TDMA mechanism to control the upstream traffic so that data packets from multiple ONUs do not collide in upstream transmission. However, the mechanism requires QoS management in an ODN network. The management cannot be implemented or may cause severe efficiency decrease because ODN is a passive network.

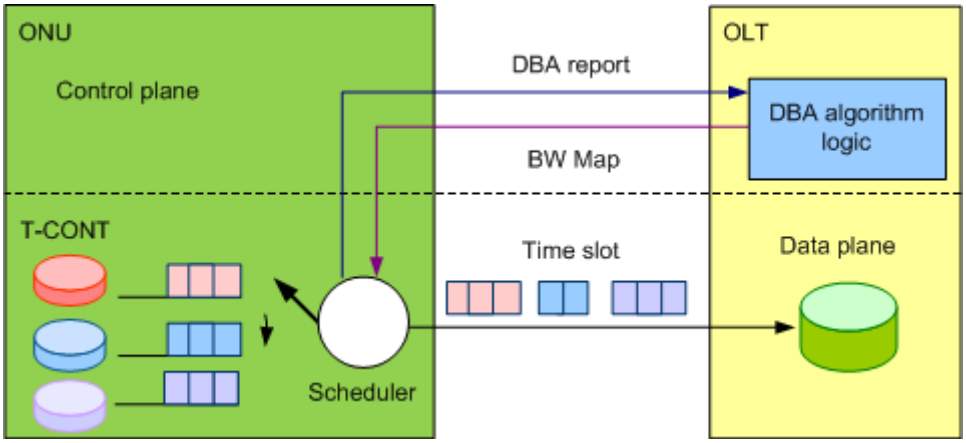
A mechanism for upstream GPON traffic management has been a primary focus in standardization of GPON traffic management. To resolve the problem, ITU-T Recommendation G.984.3 is developed, which defines the DBA protocol for managing upstream PON traffic.

DBA enables the OLT to monitor congestion on the PON network in real time. Then, the OLT can dynamically adjust bandwidths based on congestion, bandwidth usages, and configurations. DBA supports the following functions:

- Improves upstream bandwidth usages on a PON port.
- Supports more users on a PON port.
- Provides higher bandwidths for users, especially the services with significant bandwidth bursts.

Figure 2-18 shows DBA principles.

Figure 2-18 DBA principles



- The embedded DBA module of an OLT continuously collects DBA reports, performs calculation, and uses the BWMap field in the downstream frame to notify the ONU of the DBA calculation result.
- According to the BWMap information, the ONUs send data upstream in the timeslots allocated to them, and occupy the upstream bandwidth. Therefore, each ONU dynamically adjusts its upstream bandwidth according to its actually transmitted data traffic, improving upstream bandwidth usage.

Bandwidth can also be allocated in static mode, or fixed mode. In this mode, an OLT periodically allocates a fixed bandwidth to each ONU based on the ONU's service level agreement (SLA), bandwidth, and delay indicators.

- In fixed mode, an OLT uses a polling mechanism. The bandwidths allocated to ONUs may vary but the bandwidth allocated to each ONU is the same in each polling period. The bandwidth guarantee depends on an ONU's SLA but not on its upstream service traffic. An ONU is allocated a fixed bandwidth even carrying no upstream services.
- The allocation mode is simple and applies to services, such as TDM, that have a fixed traffic, but does not apply to IP services that have burst requirements on bandwidth. If the mode applies to the IP services, the upstream bandwidth usage is low because the upstream bandwidth cannot be adjusted dynamically based on the upstream service traffic.

2.7.4 FEC

In actual applications, the transmission of digital signals introduces bit errors and jitter, which degrade signal transmission quality.

To resolve the preceding issue, an error correction technology is required. Among the error correction technologies, the effective ones achieve transmission reliability by reducing bandwidth usages, which also increases telecom device complexity. The error correction technologies are used for controlling errors. The codes involved in these technologies are classified as error detection codes and error correction codes based on usage scenarios.

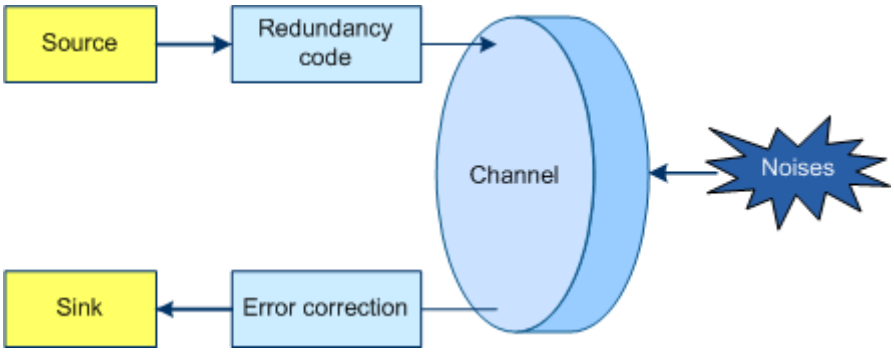
- Error detection codes, such as parity check codes, are used for detecting error codes.
- Error correction codes, such as BCH codes, Reed-Solomon (RS) codes, and Hamming codes, are used for automatically correcting errors.

The only difference between the error detection codes and error correction codes lies in performance parameters applied in different usage scenarios. FEC uses error correction codes.

FEC is a data coding technology, which enables the RX end to check error bits in transmission based on the coding data. FEC is unidirectional, not supporting error information feedback.

Redundant codes are added to signals on the TX end. Then, the RX end checks the signals for errors based on error-correcting code (ECC) and corrects errors is there is any. Common FEC codes include Hamming codes, RS codes, and convolutional codes. Figure 2-19 shows FEC principles.

Figure 2-19 FEC principles



In the GPON FEC algorithm, the most common RS code RS (255,239) is used, where the code word is 255 bytes long, consisting of 239 data bytes followed by 16 overhead redundant bytes. RS code RS (255,239) complies with ITU-T Recommendation G.984.3. The FEC algorithm drops the bit error rate (BER) of 10-3 to 10-12 for GPON lines. However, due to the overhead caused by multi-frame tail fragments, the bandwidth throughput of the GPON system with FEC enabled is about 90% of that with FEC disabled.

FEC characteristics are as follows:

- Does not require data retransmission, thereby improving real-time efficiency.
- Enables lines to tolerate louder noises on a basis of a higher bandwidth overhead. (In this case, users must balance between the transmission quality and the bandwidth usage based on site requirements.)

Based on the preceding characteristics, FEC applies to:

- The services requiring error detection and correction at the RX end without retransmission.
- Data transmission if the network is in a poor condition. For example, the transmission distance from the OLT to an ONT is long or the transmission line is of poor quality, which results in insufficient optical power budget or high BERs.
- The services requiring no delays (a retransmission prolongs the delay).

FEC status can be configured in GPON systems based on GPON ports in the downstream direction (by running the **port fec** command) and based on ONUs in the upstream direction. To configure the FEC status in the upstream direction based on ONUs, run either of the following commands:

- In profile mode, run the **fec-upstream** command.
- In discrete mode, run the **ont fec-upstream** command.

2.7.5 Line Encryption

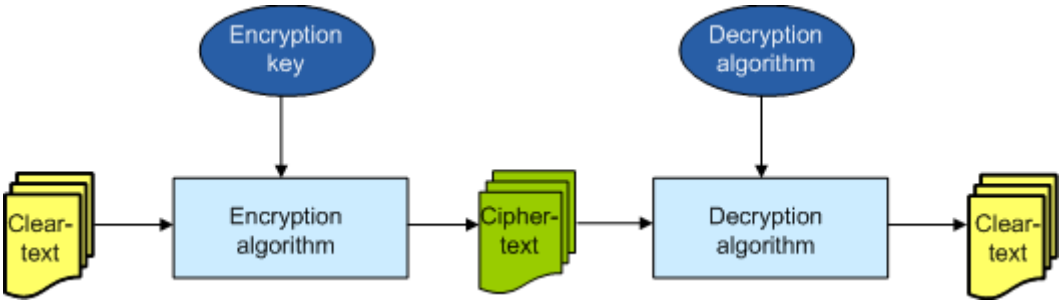
In a GPON system, downstream data is broadcast to all ONUs. Then, unauthorized ONUs can receive the downstream data of authorized ONUs, causing system risks.

Line encryption is used to eliminate these security risks. The GPON system uses the Advanced Encryption Standard 128 (AES128) algorithm to encrypt the data packets transmitted in plaintext mode so that the packets are transmitted in ciphertext mode, improving system security. Enable line encryption if the usage scenarios promote high security requirements.

- The line encryption algorithms used in GPON systems neither increase overhead nor decrease bandwidth usages.
- The line encryption algorithms will not prolong transmission delays.

Figure 2-20 shows line encryption process.

Figure 2-20 Line encryption process



Key Exchange and Switchover

1. The OLT initiates a key exchange request to the ONU. The ONU responds to the request and sends a new key to the OLT.
2. After receiving the new key, the OLT switches the key to the new one and uses the new key to encrypt data.
3. The OLT sends the frame number that uses the new key to the ONU.
4. The ONU receives the frame number and switches the verification key on data frames.

NOTE

- Due to length limitation on PLOAM messages, the ONU sends the key to the OLT in two pieces and sends both parts of the key three times for extra redundancy. If the OLT is unsuccessful in receiving either part of the key all three times it is transmitted, the OLT initiates a key exchange request to the ONU again until the OLT receives the same key for three times.
- The OLT issues a command three times to the ONU to notify the ONU of using the frame number of the new key. The ONU switches the verification key on data frames after receiving the command only once.

Configuration Method

- In GPON systems, run either of the following commands to configure line encryption status based on GEM ports (excluding multicast and broadcast GEM ports).
  - In profile mode, run the **gem add** command.
  - In discrete mode, run the **gemport add** command.

- Run either of the following commands to encrypt a GEM port in ONT management and control channels (OMCCs):
  - In profile mode, run the **omcc encrypt** command.
  - In discrete mode, run the **ont omcc encrypt** command.

2.7.6 Energy Conservation

Energy conversion enables the OLT to periodically shut down an ONU optical module when the ONU is idle, thereby conserving energy for GPON lines.

Overview

An ONU optical module is still working when the ONU is idle. Idle indicates that the traffic within the detection period is less than the specified threshold. In such a case, the OLT can periodically shuts down the ONU optical module so that it does not transmit or receive data any more. This configuration reduces ONU power consumption and conserves energy.

NOTE
Energy conservation is recommended to apply in FTTH scenario, which is more effective than in other scenarios.

Principles

Both ITU-T G.987.3 and G.984.3 define the following energy conservation modes: doze, cyclic sleep, and watchful sleep.

NOTE
Huawei OLTs support only the doze mode.

Doze Implementation

- After an ONU enters doze mode, the OLT shuts down the ONU optical module in the TX direction. In such a case, the ONU can only receive downstream data from the OLT.
- If the ONU requires to send data upstream, or the OLT requires the ONU to exit the mode for some reasons, such as for upgrading the ONU, the OLT uses a wakeup event to awake the ONU so that the ONU optical module restores in the TX direction.

Configuring Energy Conservation

1. Run the **ont power-reduction-profile add** command to create a GPON ONU energy conservation profile.

NOTE
An OLT supports up to 32 energy conservation profiles.

2. Run the **display ont power-reduction-profile** command to query the configured profile.
3. Run the **ont power-reduction-config** command to bind the energy conservation profile to the ONUs connected to a GPON port.

   After the binding, energy conservation configurations are automatically issued to the ONUs.

NOTE
- ONU energy conservation is incompatible with Type B single-homing, Type B dual-homing, Type C single homing, and Type C dual-homing. Although both ONU energy conservation and Type X homing can be configured, ONU energy conservation fails to take effect.
- Energy conservation takes effect only between Huawei OLTs and ONUs but not third-party ONUs.

- For a GPON port, the ONUs connected to it support the binding of up to eight energy conservation profiles. For an XG-PON port, the number is 16.

Key Techniques

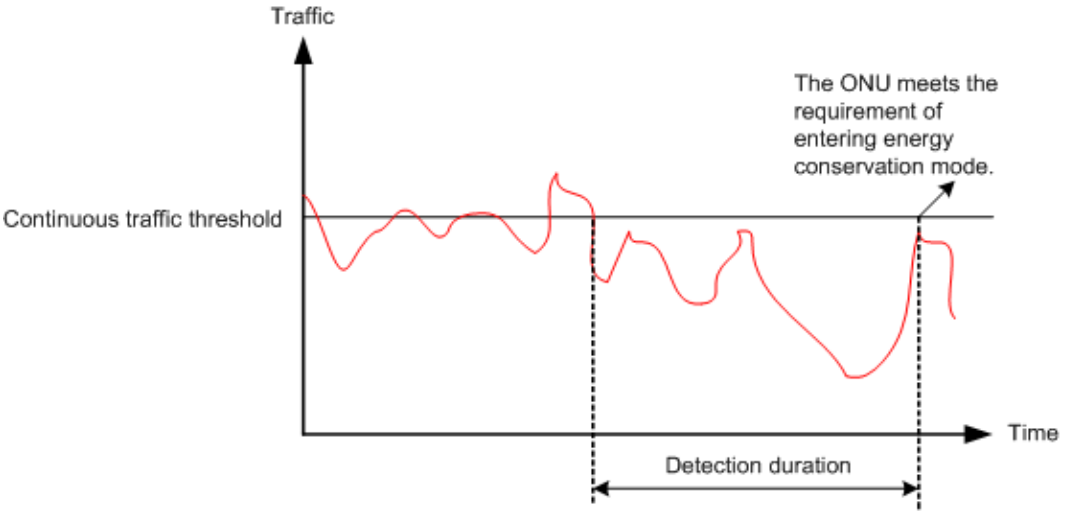
After an energy conservation profile is bound to an ONU, the ONU automatically enters or exits the doze energy conservation mode if the in-mode or out-of-mode conditions are met.

NOTE
The doze mode takes effect only in the ONU TX direction. Therefore, in-mode and out-of-mode are dedicated for upstream traffic.

In-Mode Conditions

The upstream traffic of an ONU is less than the preset continuous traffic threshold within the detection period, as shown in Figure 2-21.

Figure 2-21 In-mode conditions



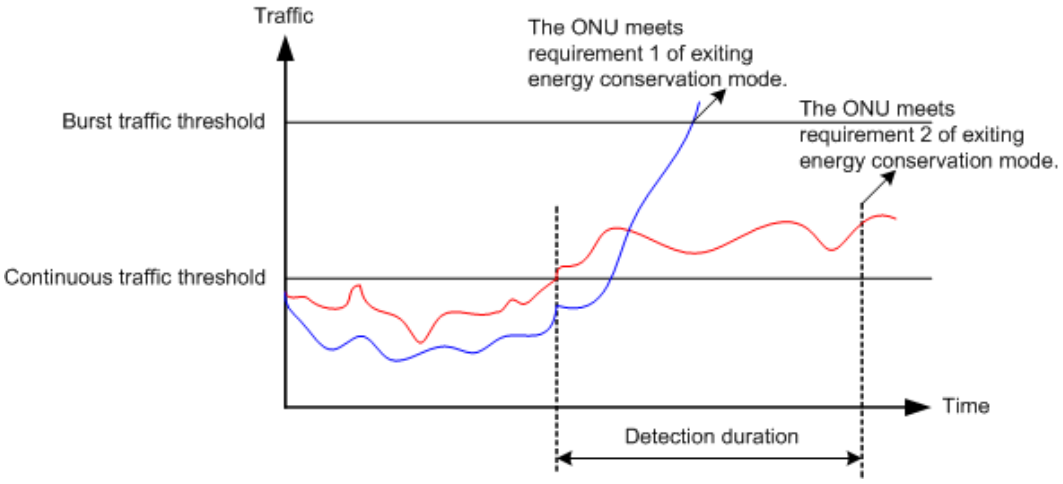
Out-of-Mode Conditions

The upstream traffic of an ONU meets either of the following requirements:

- Requirement 1: The burst traffic is greater than the configured burst traffic threshold.
- Requirement 2: The upstream traffic of the ONU is consistently greater than the continuous traffic threshold within the detection period.

Figure 2-22 shows out-of-mode conditions.

Figure 2-22 Out-of-mode conditions



2.8 GPON Networking Protection

GPON network stability is widely concerned by carriers. The line protection solutions supported by GPON networks include single homing and dual homing of both GPON type B and GPON type C.

2.8.1 GPON Type B Protection

GPON type B protection allows dual-channel redundancy protection for OLT PON ports and backbone fibers on a GPON network. This feature improves ODN network reliability and ensures service continuity.

Introduction to GPON Type B Protection

Service reliability enhancement for enterprise users and mobile users becomes a focus of carriers on PON networks. ITU-T G.984.1 defines four dual PON protection configurations, among which type B and type C are feasible. Compared with type C, type B requires a lower cost. Type B provides redundancy for the OLT, OLT PON ports, and backbone fiber. When a fault occurs on an OLT PON port or backbone fiber, services can be automatically switched to the functional optical fiber.

GPON type B applies to single-homing or dual-homing scenarios.

Figure 2-23 shows a single-homed GPON type B protection network. The protection covers the active and standby PON ports on the OLT, and the active and standby optical fibers.

Figure 2-23 Single-homed GPON type B protection network

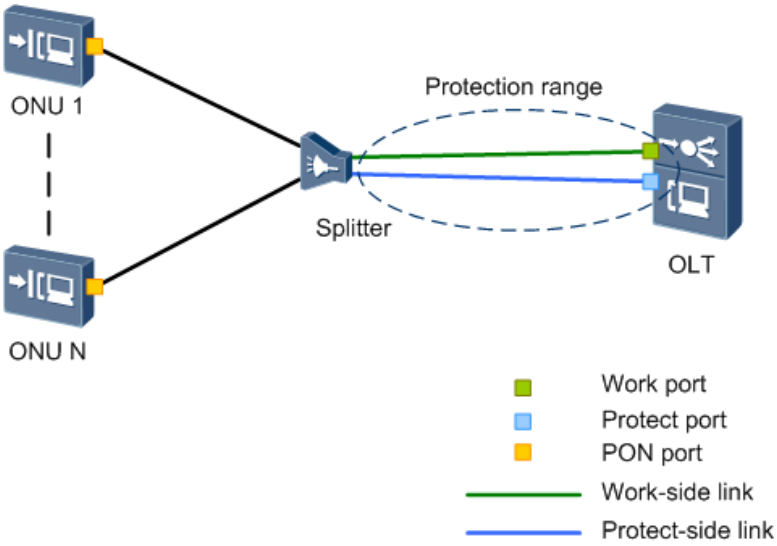
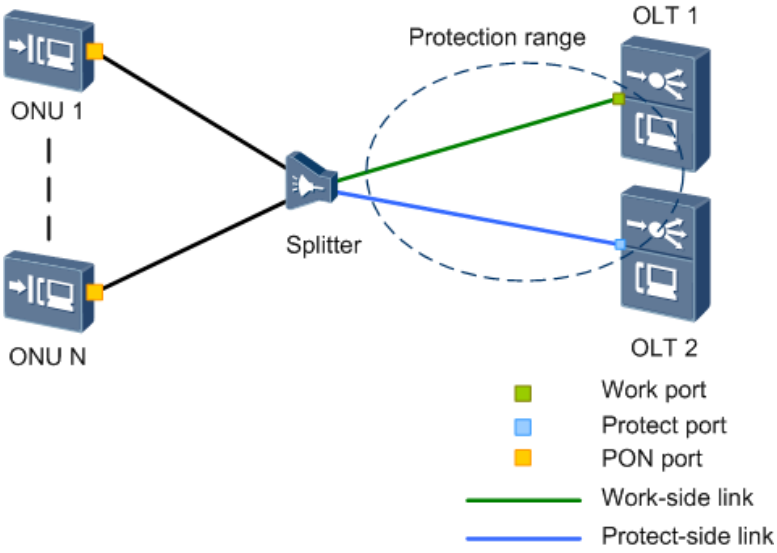


Figure 2-24 shows a dual-homed GPON type B protection network. The protection covers the active and standby OLTs, active and standby PON ports on the OLT, and the active and standby optical fibers.

Figure 2-24 Dual-homed GPON type B protection network



Networking Scenario	Advantage	Disadvantage	Usage Scenario
Single homing	Networking, OLT/ONU management, and service provisioning are	An OLT fault will interrupt services. In addition, two optical fibers routed in one pipe may both be	Protects important services, such as the enterprise private line service and the base station private line

Networking Scenario	Advantage	Disadvantage	Usage Scenario
	simple.	broken.	service.
Dual homing	Each of the two OLTs connects to a backbone fiber for remote disaster recovery.	The networking is complex, and networking costs are high. In addition, the OLT configuration is complex.	Protects important services, such as the enterprise private line service and the base station private line service. This type of networking is especially used for remote disaster recovery.

Basic Concepts of GPON Type B Protection

Protection Group

On a single-homed network, two PON ports on an OLT are added to a protection group.

NOTE

The OLT PON ports can be on the same board or on different boards. The differences are as follows:

- Port redundancy backup on the same board can conserve hardware resources. If the PON service board fails, the services on the entire board are interrupted.
- Port redundancy backup on the different boards requires hardware costs than that on the same board. If the active PON service board fails, the services can be automatically switched over to the PON ports on the standby board without being interrupted.

On a dual-homed network, the PON ports on two OLTs are added to a protection group.

Roles of Protection Group Members

Protection group members have two roles: working and protection. One protection group contains a working port and a protection port. The working port and protection port are two different PON ports. In normal cases, the working port carries services. When the link of the working port becomes faulty, the system automatically switches services from the working port to the protection port to ensure service continuity.

State of Protection Group Members

Protection group members have two states: active and standby. The active port forwards data and the standby port does not forward data.

Switching Types

The switching can be triggered automatically by a fault or performed manually. Manual operations that may cause switching are forcible switching and locking.

- In automatic switching, the OLT and ONU automatically switches to the standby link when the conditions for triggering the switching are met.

- In forcible switching, users run the **force-switch** command on the OLT to perform the switching regardless of whether specific group members are running properly.
- After switching, the working port's state becomes standby. Then, if users run the **lockout** command on the OLT to lock a group member (for a single-homed network, both working and protection ports can be locked; for a dual-homed network, only the protection port can be locked), the switching is performed and the working port's state becomes active.

Protection group members are switched only when the following conditions are met:

- The protection group is enabled.

NOTE

The status of a protection group can be queried using the **display protect-group** command on the OLT. If **Admin State** is displayed in the output, the protection group is enabled.

- The protection group is not frozen using the **freeze** command on the OLT.
- The protection group is not locked using the **lockout** command on the OLT.
- The protection group member is not forcibly switched using the **force-switch** command on the OLT.

Operation Restriction Relationships in Protection Switching

Table 2-4 Type B single homing protection switching

Current Status				Next Operation									Remarks
Enabled	Frozen	Locked	Forcible switching	Enabled	Disabled	Frozen	Unfrozen	Locked	Unlocked	Forcible switching	Canceling forcible switching	Automatic switching	None
No	No	No	No	Supported	N/A	N/A	N/A	Supported	N/A	N/A	N/A	N/A	None
No	No	Yes	No	Supported	N/A	N/A	N/A	N/A	Supported	N/A	N/A	N/A	None
Yes	No	No	No	N/A	Supported	Supported	N/A	Supported	N/A	Supported	N/A	Supported	None
Yes	No	No	Yes	N/A	Supported	Supported	N/A	Supported	N/A	Supported	N/A	N/A	The forcible switchi

Current Status				Next Operation									Remarks
													ng status will be cleared when the protection switching is disabled or locked.
Yes	No	Yes	No	N/A	Supported	Supported	N/A	N/A	Supported	N/A	N/A	N/A	None
Yes	Yes	No	No	N/A	N/A	N/A	Supported	N/A	N/A	N/A	N/A	N/A	None
Yes	Yes	No	Yes	N/A	N/A	N/A	Supported	N/A	N/A	N/A	N/A	N/A	None
Yes	Yes	Yes	No	N/A	N/A	N/A	Supported	N/A	N/A	N/A	N/A	N/A	None

NOTE

The statuses that are not listed in the preceding table, such as disabled and forcible switching, are unavailable.

Table 2-5 Type B dual homing protection switching

Current Status	Next Operation	Remarks

Current Status			Next Operation							Remarks
Enabled	Locked	Forcible switching	Enabled	Disabled	Locked	Unlocked	Forcible switching	Canceling forcible switching	Automatic switching	None
No	No	No	Supported	N/A	Supported	N/A	N/A	N/A	N/A	Protection switching is consistently enabled on the working side, and the supported statuses are available only on the protection side.
No	Yes	No	Supported	N/A	N/A	Supported	N/A	N/A	N/A	Protection switching is consistently enabled on the working side, and the

Current Status			Next Operation							Remarks
										supported statuses are available only on the protection side.
Yes	No	No	N/A	Supported	Supported	N/A	Supported	N/A	Supported	Protection switching can be disabled or locked only on the protection side.
Yes	No	Yes	N/A	Supported	Supported	N/A	Supported	Supported	N/A	Protection switching can be disabled or locked only on the protection side. In either of the statuses, the forcib

Current Status			Next Operation							Remarks
										le switching status will be cleared.
Yes	Yes	No	N/A	Supported	N/A	Supported	N/A	N/A	N/A	Protection switching can be locked only on the protection side.

NOTE

The statuses that are not listed in the preceding table, such as disabled and forcible switching, are unavailable.

Associated Switching

Associated switching is implemented on a dual-homed network as follows: A protection group is associated on the OLT with the uplink Ethernet port status and BFD/MEP session status. In such a case, when the OLT's upper-layer network or the Layer 2 OLT physical link fails, the active OLT triggers a dual-homing protection switchover so that services will be switched to the standby OLT.

Single-Homing GPON Type B Protection Principles

On a single-homed network, the two PON ports on the OLT are in active/standby state, and they cannot forward packets at the same time. When the active link is faulty due to an optical path or PON port fault, the ONU can rapidly switch services to the standby link. An automatic switchover can be triggered by any of the following conditions:

- Active fiber cut
- Active PON port failure
- Line quality deterioration

NOTE

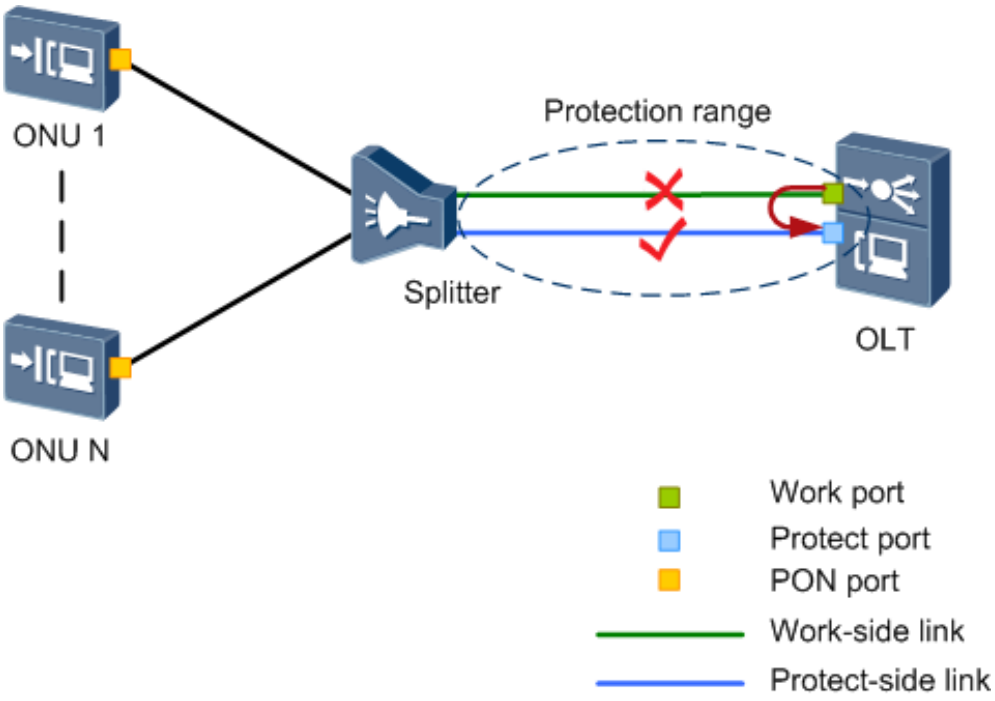
When the line quality deteriorates, and the BER reaches the preset threshold, the ONU goes offline, triggering a protection switchover. To configure a BER threshold (consisting of the failed signal of ONU threshold and degraded signal of ONU threshold), run the **gpon alarm-profile add** command.

The following section describes the GPON type B protection switching in different scenarios.

Scenario 1: Active Optical Fiber Is Cut

The active optical fiber is cut when the OLT PON port is working, as shown in Figure 2-25.

Figure 2-25 Active optical fiber is cut

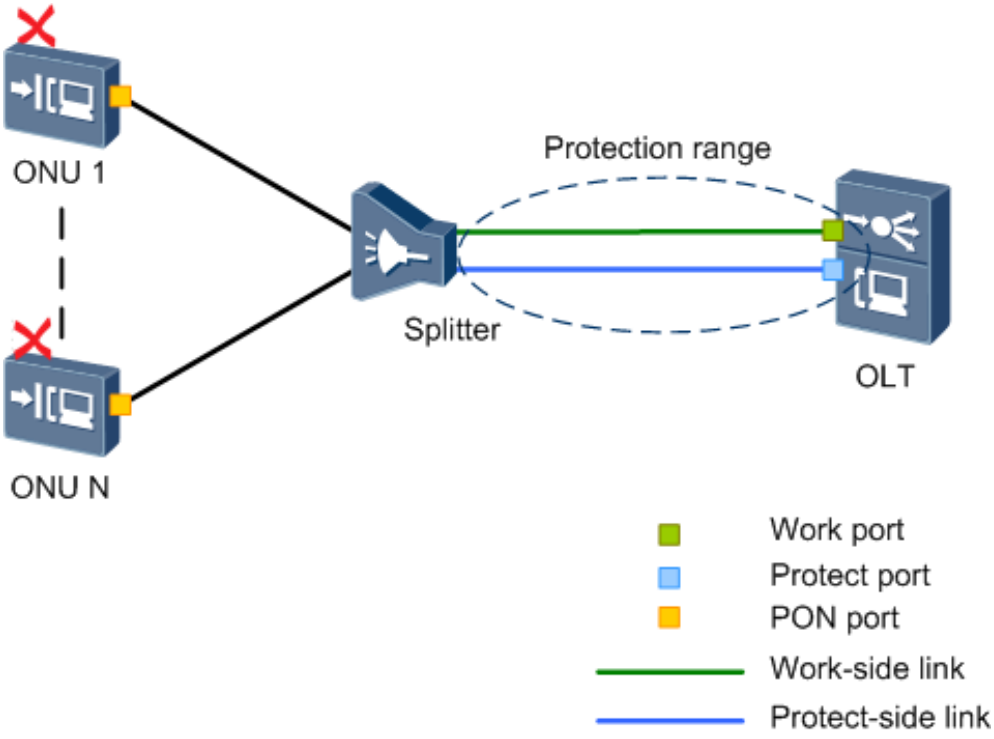


- The working port is in the active state and is working properly. The protection port is in the standby state.
- When detecting a loss of signal (LOS) alarm (generated due to the active optical fiber cut), the working port disables the transmission of the optical module.
- When detecting an LOS alarm of the working port, the protection port enables the transmission of the optical module and performs ONU ranging.
- If the optical fiber connected to the protection port is functional, and ONU ranging is successful, the protection port reports an LOS clear alarm.
- The working port switches to the standby state. The protection port switches to the active state. Then, the protection switching ends.

Scenario 2: All ONUs Go Offline

All ONUs connected to the OLT PON port go offline, as shown in Figure 2-26.

Figure 2-26 All ONUs go offline



- The working port is in the active state and is working properly. The protection port is in the standby state.
- When detecting an LOS alarm (generated because all ONUs go offline), the working port disables the transmission of the optical module.
- When detecting an LOS alarm of the working port, the protection port enables the transmission of the optical module and performs ONU ranging.
- No ONU connected to a PON port goes online due a ranging failure. Therefore, the OLT cyclically detects the working and protection ports until an ONU goes online.
- After the ONU goes online, switching is performed between the PON ports if the protection port is detected. If the protection port is not detected, the working port continues working.

Dual-Homing GPON Type B Protection Principles

On a dual-homed network, two OLTs are in active/standby state, and they cannot forward packets at the same time. Users must manually configure the same service data on the two OLTs so that the ONU can rapidly switch services from the active OLT to the standby one when the active OLT becomes faulty due to an optical path or component fault. This shortens service interruption duration. An automatic switchover can be triggered by any of the following conditions:

NOTE
Currently, type B dual-homing protection mainly be used in the passive optical LAN (POL) solution and requires cooperation of the OLT, ONU, and aggregation device.

- Optical fiber cut from the active OLT
- Active OLT fault

- Active OLT PON port fault
- The OLT's uplink is faulty (this condition triggers automatic switching only in the associated protection switching scenario).
- Active line quality deterioration

NOTE

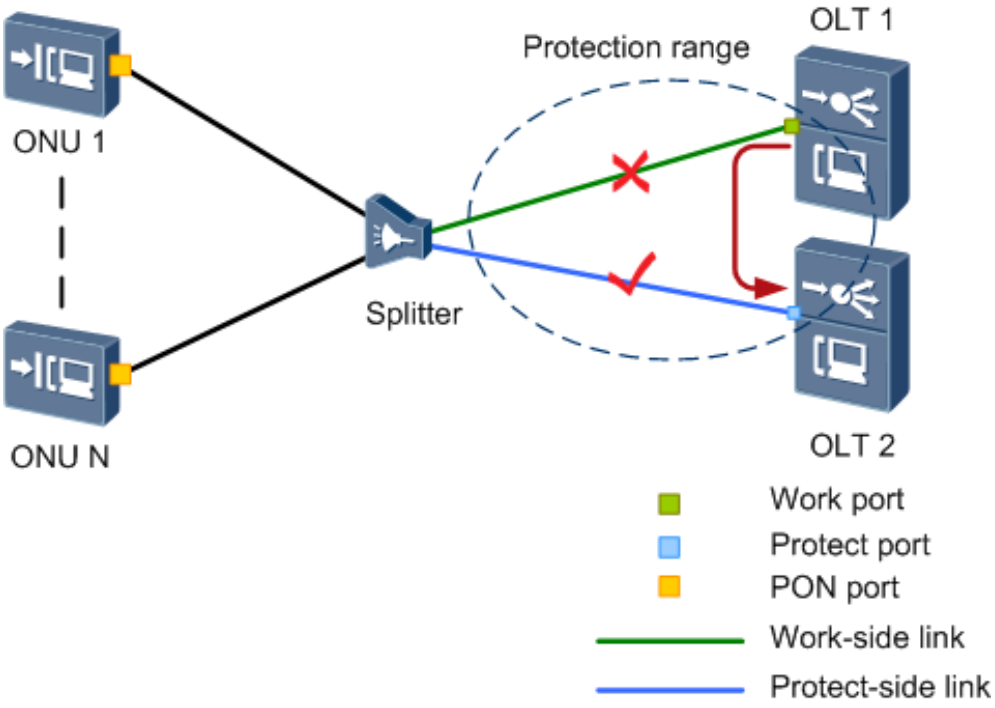
When the line quality deteriorates, and the BER reaches the preset threshold, the ONU goes offline, triggering a protection switchover. To configure a BER threshold (consisting of the failed signal of ONU threshold and degraded signal of ONU threshold), run the **gpon alarm-profile add** command.

The following section describes the GPON type B protection switching in different scenarios.

Scenario 1: Active Optical Fiber Is Cut

The active optical fiber is cut when the OLT PON port is working, as shown in Figure 2-27.

Figure 2-27 Active optical fiber is cut



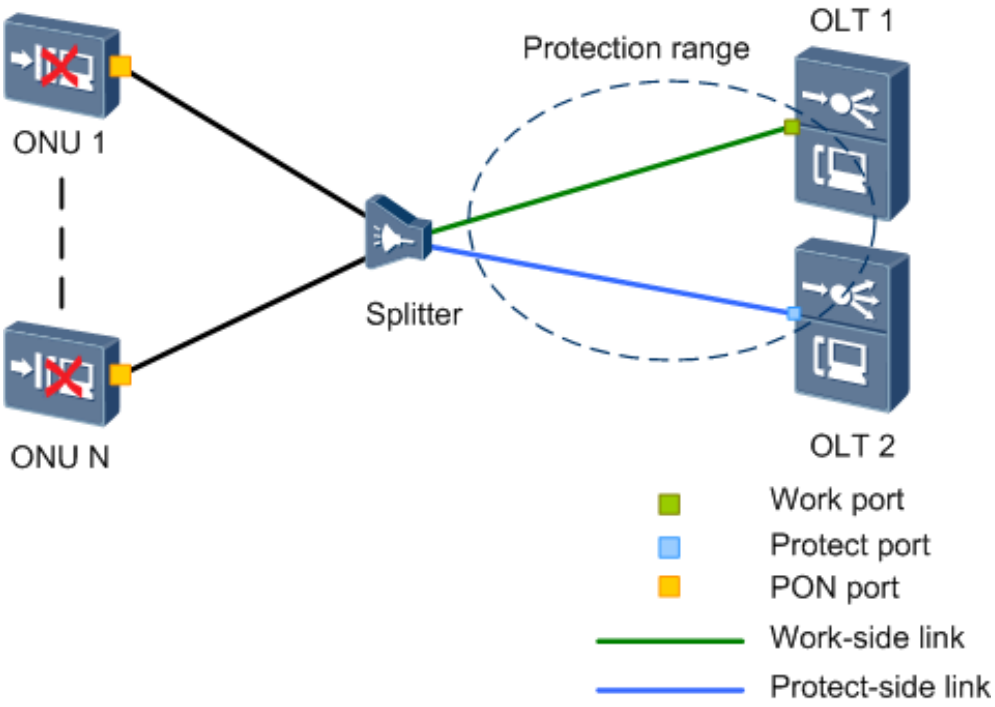
The protection switching process is as follows:

- The working port is in the active state and is working properly. The protection port is in the standby state.
- When detecting a loss of signal (LOS) alarm (generated due to the active optical fiber cut), the working port disables the transmission of the optical module.
- When detecting an LOS alarm of the working port, the protection port enables the transmission of the optical module and performs ONU ranging.
- If the optical fiber connected to the protection port is functional, and ONU ranging is successful, the protection port reports an LOS clear alarm.
- The working port switches to the standby state. The protection port switches to the active state. Then, the protection switching ends.

Scenario 2: All ONUs Go Offline

All ONUs connected to the OLT PON port go offline, as shown in Figure 2-28.

Figure 2-28 All ONUs go offline



The protection switching process is as follows:

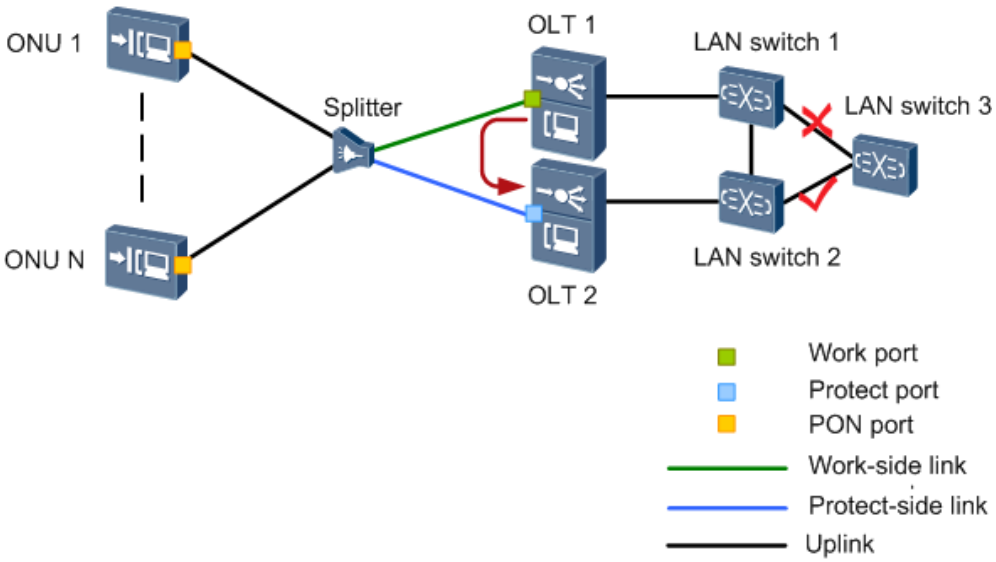
- The working port is in the active state and is working properly. The protection port is in the standby state.
- When detecting an LOS alarm (generated because all ONUs go offline), the working port disables the transmission of the optical module.
- When detecting an LOS alarm of the working port, the protection port enables the transmission of the optical module and performs ONU ranging.
- No ONU connected to a PON port goes online due a ranging failure. Therefore, the OLT cyclically detects the working and protection ports until an ONU goes online.
- After the ONU goes online, switching is performed between the PON ports if the protection port is detected. If the protection port is not detected, the working port continues working.

Scenario 3: Associated Protection Switching Is Caused by a Connection Failure in the OLT's Upstream Transmission Network

If a BFD session has been configured on the OLT, the BFD session can be bound to a protection group for creating an association between them. If CFM has been enabled on an OLT, an MEP session can be bound to a protection group for creating an association between them. Based on the associations, when the upper-layer network connection of the active OLT fails, the active and standby OLTs perform a switchover and notifies the ONU of the

switchover. In this way, services are restored. Figure 2-29 shows the associated protection switching caused by a connection failure in the OLT's upstream transmission network.

Figure 2-29 OLT's upstream transmission network connection fails



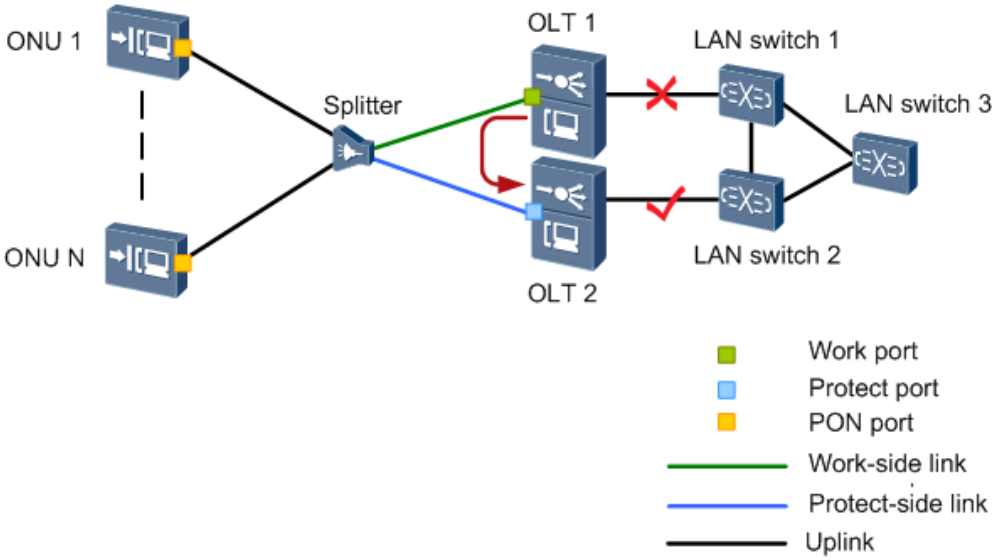
The protection switching process is as follows:

- On the OLT, the dual-homing protection group is associated with the BFD or MEP session. When the upstream route of the working OLT (OLT 1) fails, OLT 1 checks whether the protection OLT (OLT 2) and the upstream route of OLT 2 are functional. If they are functional, and the both OLTs do not carry out a forcible switchover or locking operation, OLT 1 data has been synchronized to OLT 2. Then, the two OLTs perform a switchover.
- OLT 2 starts to work. It enables the transmission of the optical module and performs ONU ranging.
- After the switchover, the ONU service data is sent to OLT 2 through the protection port, and service data is transmitted over the protection link.
  - OLT 1 becomes standby.
  - OLT 2 becomes active.

Scenario 4: Associated Protection Switching Is Caused by an OLT's Physical Link Fault

An OLT protection group is associated with the uplink Ethernet port status. Based on the association, when the physical link of the active OLT fails, the two OLTs perform a switchover, and the active OLT notifies the ONU of the switchover. In this way, services are restored. Figure 2-30 shows the associated protection switching caused by an OLT's physical link fault.

Figure 2-30 OLT's physical link fails



The protection switching process is as follows:

- On the OLT, the dual-homing protection group is associated with the uplink Ethernet port status. When the Ethernet port associated with the protection group on the working OLT (OLT 1) becomes **Down**, OLT 1 checks whether the protection OLT (OLT 2) and the physical link of the protection OLT are functional. If they are functional, and the both OLTs do not carry out a forcible switchover or locking operation, OLT 1 data has been synchronized to OLT 2. Then, the two OLTs perform a switchover.
- OLT 2 starts to work. It enables the transmission of the optical module and performs ONU ranging.
- After the switchover, the ONU service data is sent to OLT 2 through the protection port, and service data is transmitted over the protection link.
  - OLT 1 becomes standby.
  - OLT 2 becomes active.

2.8.2 GPON Type C Protection

The GPON type C protection switching is implemented through the redundancy configuration of the two PON ports on the ONU, backbone optical fiber, optical splitter, and tributary optical fiber on a GPON network. Each item is in a dual configuration. The protection improves the reliability on the optical distribution network (ODN) and prevents service interruption.

Introduction to GPON Type C Protection

Service reliability enhancement for enterprise users and mobile users becomes a focus of carriers on passive optical network (PON) networks. G.984.1 (approved in 2008) defines four dual PON protection configurations, among which type B and type C are feasible. Compared with type B, type C provides higher reliability. Type C provides redundancy for OLT (dual homing), ONU's PON ports, backbone fibers, optical splitters, and distribution optical fibers. When a fault occurs, services can be automatically switched to the functional link. After the fault is rectified, services are automatically switched back to the original link.

GPON type C protection can be deployed in two networking scenarios: single homing and dual homing.

Figure 2-31 shows the GPON type C protection single homing network.

Figure 2-31 GPON type C protection network (single homing)

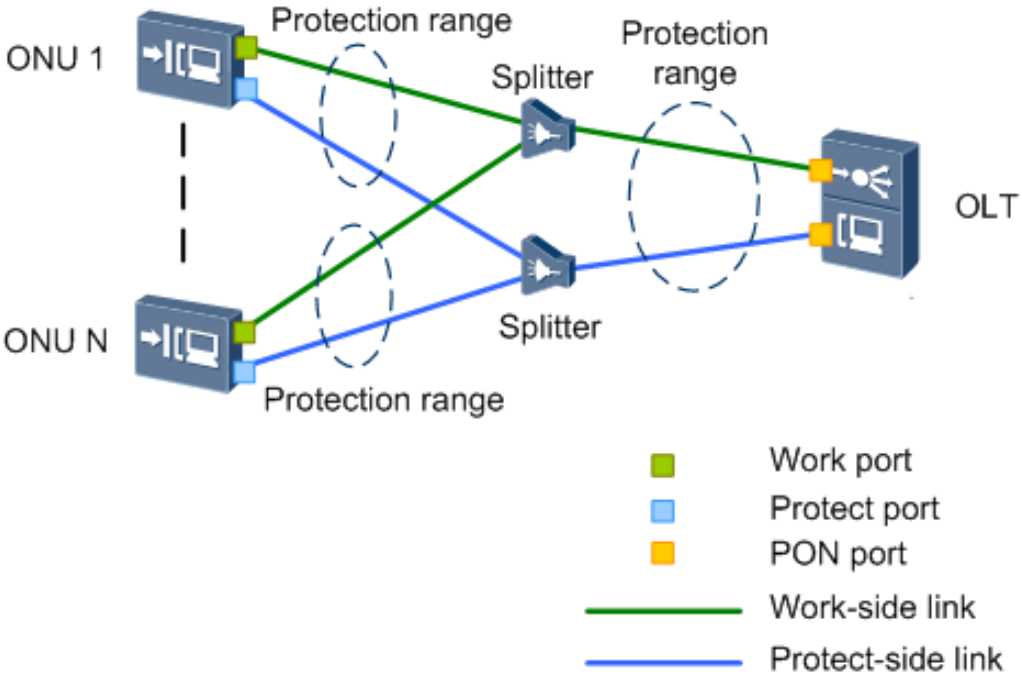
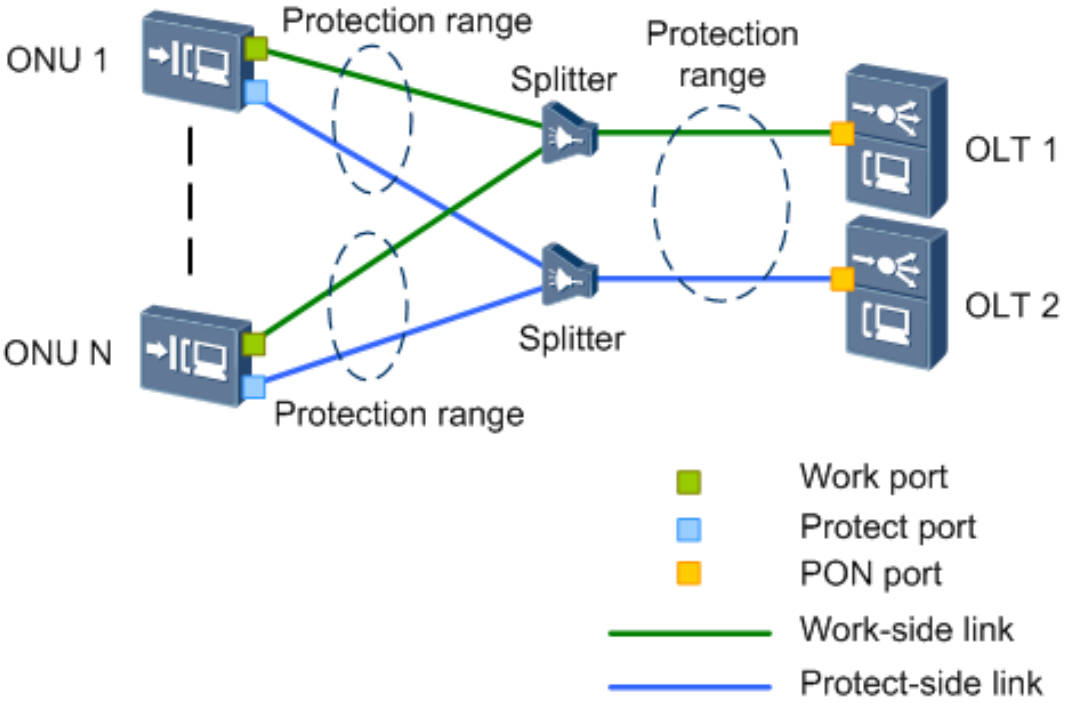


Figure 2-32 shows the GPON type C protection dual homing network.

Figure 2-32 GPON type C protection network (dual homing)



Networking Mode	Advantage	Disadvantage	Scenario
Single homing	The networking mode is simple, and OLT and ONU can be managed easily.	When the OLT becomes faulty, services are interrupted. Optical fibers are deployed on the same channel and therefore two optical fibers may be broken at the same time.	This mode is used to protect important services, such as Enterprise private line services and base station services.
Dual homing	When the active OLT or its uplink fails, services can be switched to the standby OLT.	The networking mode is complicated and costly, and the ONU management is difficult.	This mode is used to protect a power system or Enterprise private line services and base station services.

Basic Concepts of GPON Type C Protection

Protection Group

On a single-homed network, two PON uplink ports on an ONU connected to different PON ports on an OLT are added to a protection group using the CLI or NMS.

NOTE

The OLT PON ports can be on the same board or on different boards. The differences are as follows:

- Port redundancy backup on the same board can conserve hardware resources. If the PON service board fails, the services on the entire board are interrupted.
- Port redundancy backup on the different boards requires hardware costs than that on the same board. If the active PON service board fails, the services can be automatically switched over to the PON ports on the standby board without being interrupted.

On a dual-homed network, two PON uplink ports on an ONU connected to two OLTs are added to a protection group using the CLI or NMS. Switching can be performed between two members in a protection group.

Roles of Protection Group Members

Protection group members have two roles: working and protection. One protection group contains a working port (the member's role is working) and a protection port (the member's role is protection). The working port and protection port are two different uplink PON ports on the ONU. In normal cases, the working port carries services. When the link of the working port becomes faulty, the system automatically switches services from the working port to the protection port to ensure service continuity.

State of Protection Group Members

Protection group members have two states: active and standby. The active port forwards data and the standby port does not forward data.

Switching Types

The switching can be triggered automatically by a fault or performed manually. Operations that may cause switching are locking, forcible switching, and manual switching.

- In automatic switching, the OLT and ONU automatically switches to the standby link when the conditions for triggering the switching are met.
- In manual switching, users manually switch the protection group by running the **manual-switch** command on the OLT.
- In forcible switching, users run the **force-switch** command on the OLT to perform the switching regardless of whether specific group members are running properly.
- After switching, the working port's state becomes standby. Then, if users run the **lockout** command on the OLT to lock a group member (only the protection port can be locked), the switching is performed and the working port's state becomes active.

NOTE

- In training switching, users run the **exercise-switch** command on the OLT to perform the switching to test the Automatic Protection Switching (APS) function on the ports in a protection group. Services are not switched.
- In automatic switchback, when the working member in the PG recovers to the normal state, the PG automatically switches over after the WTR time expires, and service is still carried on the working member.

Protection group members are switched only when the following conditions are met:

- The protection group is enabled.

NOTE

The status of a protection group can be queried using the **display protect-group** command on the OLT. If **Admin State** is displayed in the output, the protection group is enabled.

- The protection group is not locked using the **lockout** command on the OLT.
- The protection group member is not forcibly switched using the **force-switch** command on the OLT.

Operation Restriction Relationships in Protection Switching

Table 2-6 Type C single homing protection switching

Current Status			Next Operation									Remarks
Enabled	Locked	Forcible switching	Enabled	Disabled	Locked	Unlocked	Forcible switching	Canceling forcible switching	Automatic switching	Manual switching	Training switching	None
No	No	No	Supported	N/A	Supported	N/A	N/A	N/A	N/A	N/A	N/A	None
No	Yes	No	Supported	N/A	N/A	Supported	N/A	N/A	N/A	N/A	N/A	None
Yes	No	No	Not supported	Supported	Supported	N/A	Supported	N/A	Supported	Supported	Supported	None
Yes	No	Yes	Not supported	Supported	Supported	N/A	Supported	Supported	N/A	N/A	N/A	The forcible switching status will be cleared when the protection switching is disa

Current Status			Next Operation									Remarks
												bled or locked.
Yes	Yes	No	Not supported	Supported	N/A	Supported	N/A	N/A	N/A	N/A	N/A	None

NOTE

The statuses that are not listed in the preceding table, such as disabled and forcible switching, are unavailable.

Table 2-7 Type C dual homing protection switching

Current Status			Next Operation									Remarks
Enabled	Locked	Forcible switching	Enabled	Disabled	Locked	Unlocked	Forcible switching	Canceling forcible switching	Automatic switching	Manual switching	Training switching	None
No	No	No	Supported	N/A	Supported	N/A	N/A	N/A	N/A	N/A	N/A	Protection switching is consistently enabled on the working side, and the supported statu

Current Status			Next Operation									Remarks
												ses are available only on the protection side.
No	Yes	No	Supported	N/A	N/A	Supported	N/A	N/A	N/A	N/A	N/A	Protection switching is consistently enabled on the working side, and the supported statuses are available only on the protection side.
Yes	No	No	Not supported	Supported	Supported	N/A	Supported	N/A	Supported	Supported	Supported	Protection swit

Current Status			Next Operation									Remarks
												ching can be disabled or locked only on the protection side.
Yes	No	Yes	Not supported	Supported	Supported	N/A	Supported	Supported	N/A	N/A	N/A	Protection switching can be disabled or locked only on the protection side. In either of the statuses, the forcible switching status

Current Status			Next Operation									Remarks
												will be cleared.
Yes	Yes	No	Not supported	Supported	N/A	Supported	N/A	N/A	N/A	N/A	N/A	Protection switching can be locked only on the protection side.

NOTE

The statuses that are not listed in the preceding table, such as disabled and forcible switching, are unavailable.

Associated Protection Switching

Associated switching is implemented on a dual-homed network as follows: A protection group is associated on the OLT with the uplink Ethernet port status and BFD/MEP session status. In such a case, when the OLT's upper-layer network or the Layer 2 OLT physical link fails, the OLT determines a protection switchover, ensuring service continuity. Associated protection switching applies on the network enabled with automatic site information transmission.

Single-Homing GPON Type C Protection Principles

On a single homing network, one ONU is connected to two ports on an OLT, one working as the active port and the other as standby. The two ports on the OLT cannot forward packets at the same time. When the active link is interrupted due to an optical fiber or component fault, the ONU quickly switches services to the standby PON port on the OLT. An automatic switching can be triggered by any of the following conditions:

- Loss of signal (LOS) occurs in the input direction.
- The ONU is offline.
- The OLT or ONU hardware is faulty.

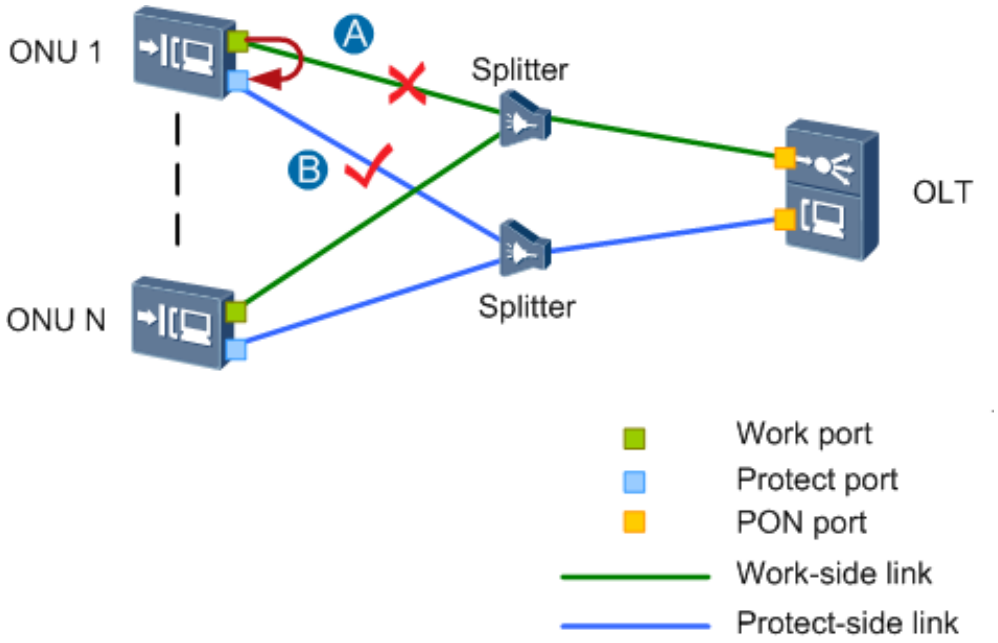
The following section describes PON switching processes in three scenarios. The OLT's PON ports are running properly, and the ONU has registered with the OLT. Users have issued the

same settings of the PON line to PON port 1 and PON port 2 on the OLT so that services can recover after protection switching.

Scenario 1: Branch Fiber Connected to a Single ONU Becomes Faulty

Figure 2-33 shows the scenario in which the branch fiber connected to a single ONU becomes faulty.

Figure 2-33 Branch fiber connected to a single ONU becomes faulty



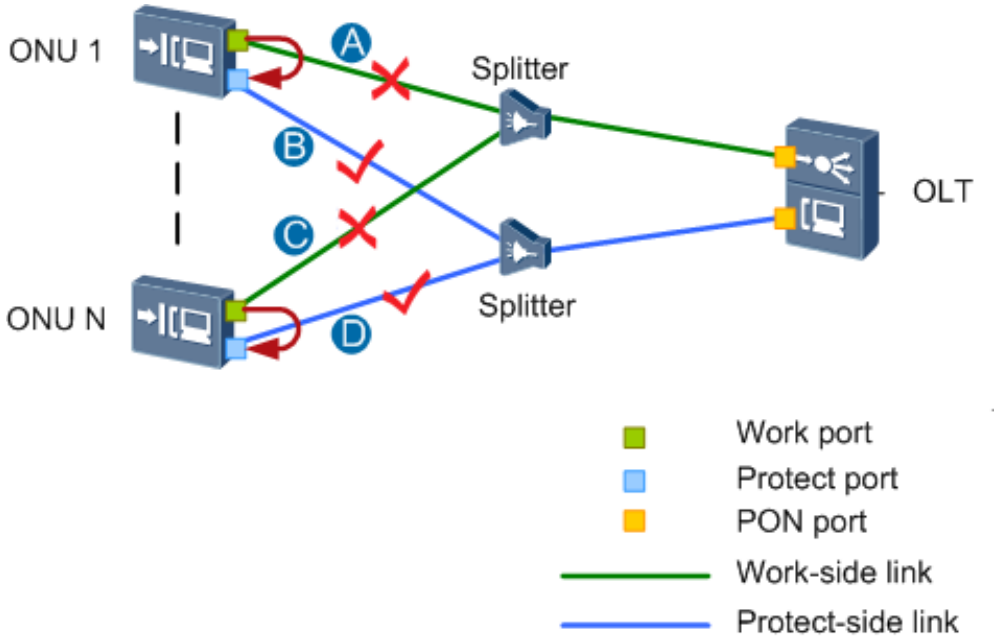
- Both the ONU and OLT check the link status and determine whether to trigger protection switching based on the link status.
  - If the OLT detects a fault on link A in Figure 2-33, it automatically switches to the protection-side link and uses the protection-side link to send messages to ONU 1 to notify that protection switching has occurred. In addition, the OLT notifies ONU 1 of the switching cause.
  - If ONU 1 detects a fault on link A in Figure 2-33, it automatically switches to the protection-side link and sends messages to the OLT to notify that protection switching has occurred. In addition, the ONU notifies the OLT of the switching cause.
- After switching, services on ONU 1 are transmitted to the OLT through the protection port, all the backbone fibers connected to the OLT transmitted service packets, and ONU N is not affected. The changes on ONU 1 are as follows:
  - The state of the working port changes to standby.
  - The state of the protection port changes to active, and service packets are transmitted through link B in Figure 2-33.
- After protection switching, ONU 1 can automatically switch back to the working port. The OLT sends an automatic switchback message and the switchback time, called the wait to restore (WTR) time, to the ONU.

- If the OLT learns that ONU 1's working port and the working-side line are functioning properly and the working-side line stays normal during the WTR time, the OLT automatically switches back to the working-side line when the WTR time expires. In addition, the OLT notifies the ONU of the switching and switching cause.
- If ONU 1 learns that its working port and the working-side line are functioning properly and the working-side line stays normal during the WTR time, the ONU automatically switches back to the working-side line when the WTR time expires. In addition, the ONU notifies the OLT of the switching and switching cause.

Scenario 2: All Branch Fibers Connected to the ONU Become Faulty

Figure 2-34 shows the scenario in which all branch fibers connected to the ONU become faulty.

Figure 2-34 All branch fibers connected to the ONU become faulty



- Both the ONU and OLT check the link status and determine whether to trigger protection switching based on the link status.
  - If the OLT detects that all branch fibers connected to the working-side line become faulty, it automatically switches to the protection-side link and uses the protection-side link to send messages to all the ONUs to notify that protection switching has occurred. In addition, the OLT notifies all the ONUs of the switching cause.
  - If the ONU detects a fault on all the branch fibers connected to the working-side link, it automatically switches to the protection-side link and sends messages to the OLT to notify that protection switching has occurred. In addition, the ONU notifies the OLT of the switching cause.
- After switching, services on the ONU are transmitted to the OLT through the protection port (that is, service packets are transmitted by the protection-side link). The changes on the ONU are as follows:

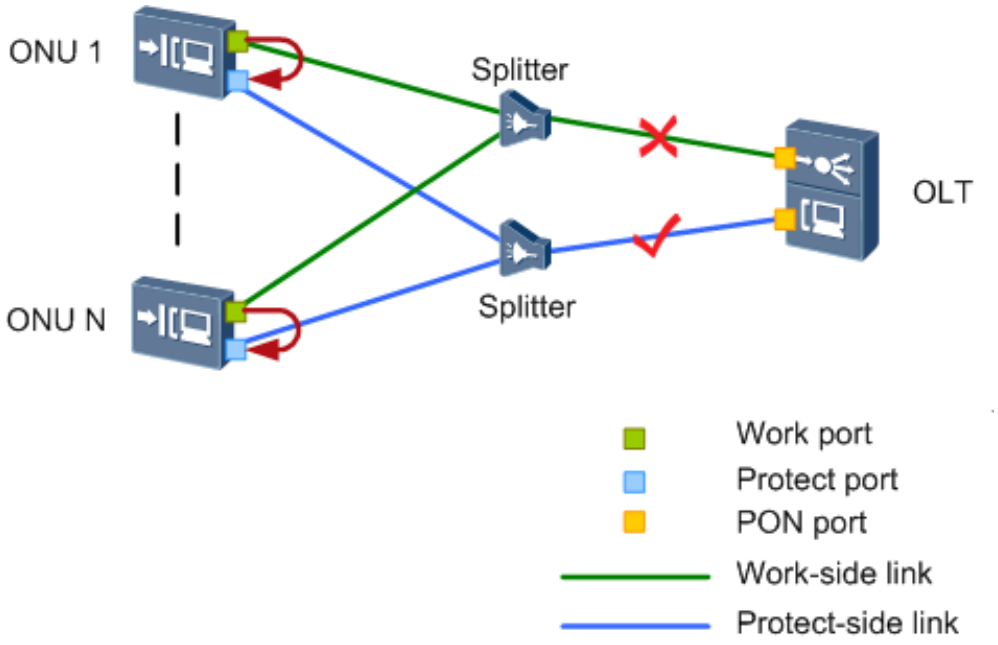
- The state of the working port changes to standby.
- The state of the protection port changes to active.

- After protection switching, the ONU can automatically switch back to the working port. The OLT sends an automatic switchback message and the switchback time, called the WTR time, to the ONU.
  - If the OLT learns that the working port and working-side links are functioning properly and link A stays normal during the WTR time, the OLT automatically switches back to the working-side link when the WTR time expires. In addition, the OLT notifies the ONU of the switching and switching cause.
  - If the ONU learns that the working port and working-side link are functioning properly and link A stays normal during the WTR time, the ONU automatically switches back to the working-side link when the WTR time expires. In addition, the ONU notifies the OLT of the switching and switching cause.

Scenario 3: Backbone Fiber Becomes Faulty

Figure 2-35 shows the scenario in which the backbone fiber becomes faulty.

Figure 2-35 Backbone fiber becomes faulty



- Both the ONU and OLT check the link status and determine whether to trigger protection switching based on the link status.
  - If the OLT detects a fault on the working-side link, it automatically switches to the protection-side link and uses the protection-side link to send messages to all the ONUs to notify that protection switching has occurred. In addition, the OLT notifies all the ONUs of the switching cause.
  - If the ONU detects a fault on the working-side link, it automatically switches to the protection-side link and sends messages to the OLT to notify that protection switching has occurred. In addition, the ONU notifies the OLT of the switching cause.

- After switching, services on the ONU are transmitted to the OLT through the protection port (that is, service packets are transmitted by the protection-link). The changes on the ONU are as follows:
  - The state of the working port changes to standby.
  - The state of the protection port changes to active.
- After protection switching, the ONU can automatically switch back to the working port. The OLT sends an automatic switchback message and the switchback time, called the WTR time, to the ONU.
  - If the OLT learns that the working port and working-side links are functioning properly and link A stays normal during the WTR time, the OLT automatically switches back to the working-side link when the WTR time expires. In addition, the OLT notifies the ONU of the switching and switching cause.
  - If the ONU learns that the working port and working-side link are functioning properly and link A stays normal during the WTR time, the ONU automatically switches back to the working-side link when the WTR time expires. In addition, the ONU notifies the OLT of the switching and switching cause.

Dual-Homing GPON Type C Protection Principles

On a dual homing network, two PON lines, one working as the active line and one as standby, between an ONU and two OLTs cannot forward packets at the same time. When the active line is interrupted due to an optical fiber or component fault, the ONU quickly switches services to the OLT connected to the standby line (called the protection OLT). An automatic switchover can be triggered by any of the following conditions:

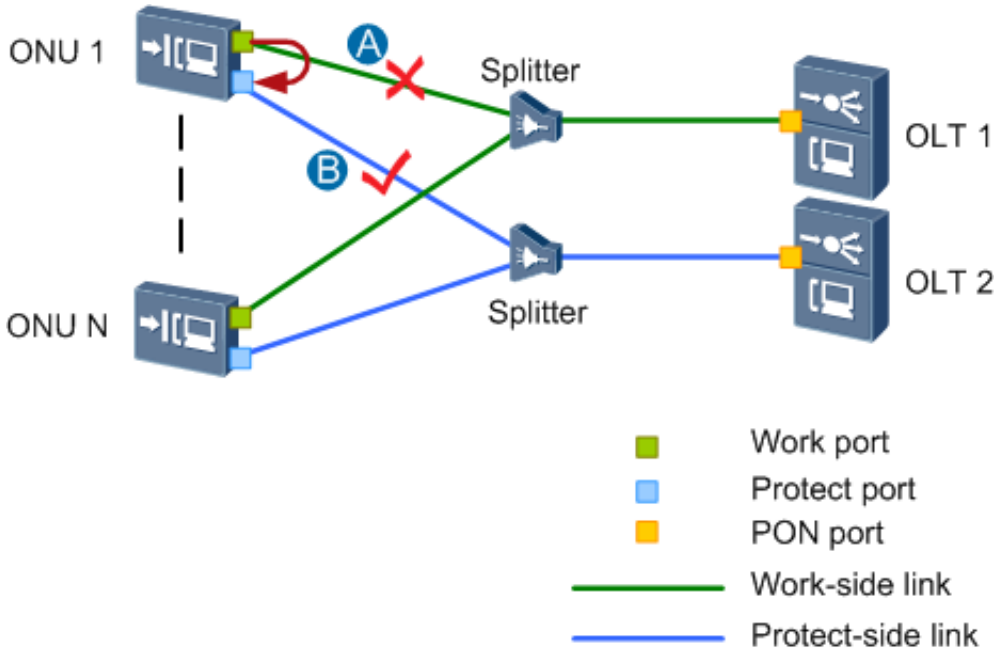
- Loss of signal (LOS) occurs in the input direction.
- The ONU is offline.
- The OLT or ONU hardware is faulty.
- The OLT's uplink is faulty (this condition triggers automatic switching only in the associated protection switching scenario).

The following section describes the PON switching processes in five scenarios. OLT 1 and OLT 2 are running properly, and the ONU has registered with the OLTs. Users issue the ONU's configurations to OLT 1 and OLT 2 so that services can recover after the switching.

Scenario 1: Branch Fiber Connected to a Single ONU Becomes Faulty

Figure 2-36 shows the scenario in which the branch fiber connected to a single ONU becomes faulty.

Figure 2-36 Branch fiber connected to a single ONU becomes faulty

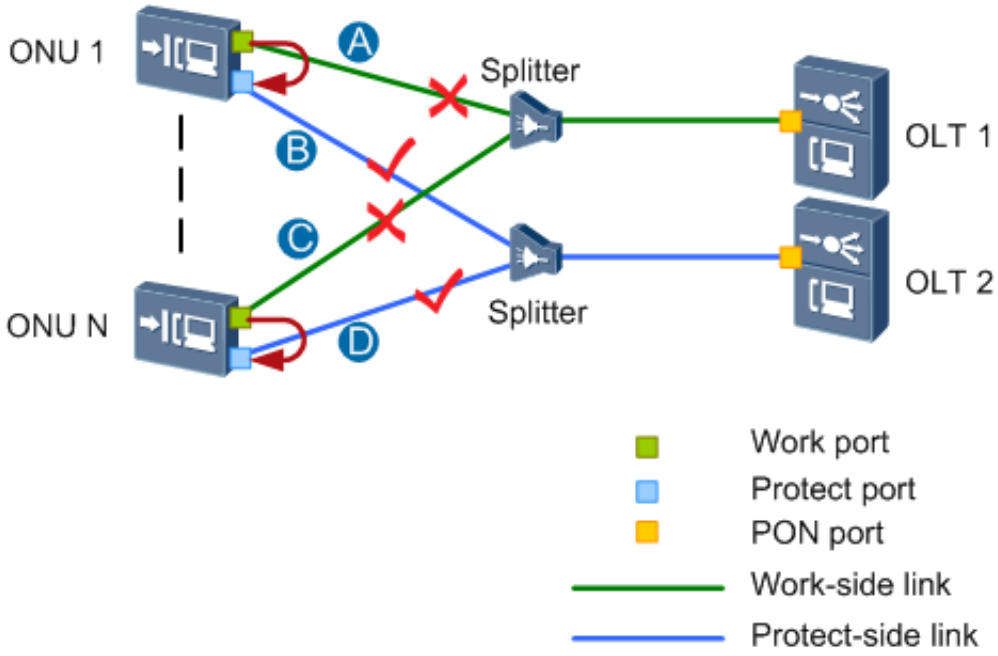


- If ONU 1 detects a fault on link A in Figure 2-36, it automatically switches to the protection-side link and sends messages to OLT 2 to notify that protection switching has occurred. In addition, the ONU notifies OLT 2 of the switching cause.
- After switching, services on ONU 1 are transmitted to the OLT through the protection port, all the backbone fibers connected to the OLT transmit service packets, and ONU N is not affected. The changes on ONU 1 are as follows:
  - The state of the working port changes to standby.
  - The state of the protection port changes to active and service packets are transmitted through link B in Figure 2-36.
- After protection switching, ONU 1 can automatically switch back to the working port. The OLT sends an automatic switchback message and the switchback time, called the wait to restore (WTR) time, to the ONU. If ONU 1 learns that the working port, working-side link, and the uplink of OLT 1 are functioning properly and link A stays normal during the WTR time, ONU 1 automatically switches to the working-side link when the WTR time expires.

Scenario 2: All Branch Fibers Connected to the Active Link Become Faulty

Figure 2-37 shows the scenario in which all branch fibers connected to the active link become faulty.

Figure 2-37 All branch fibers connected to the active link become faulty

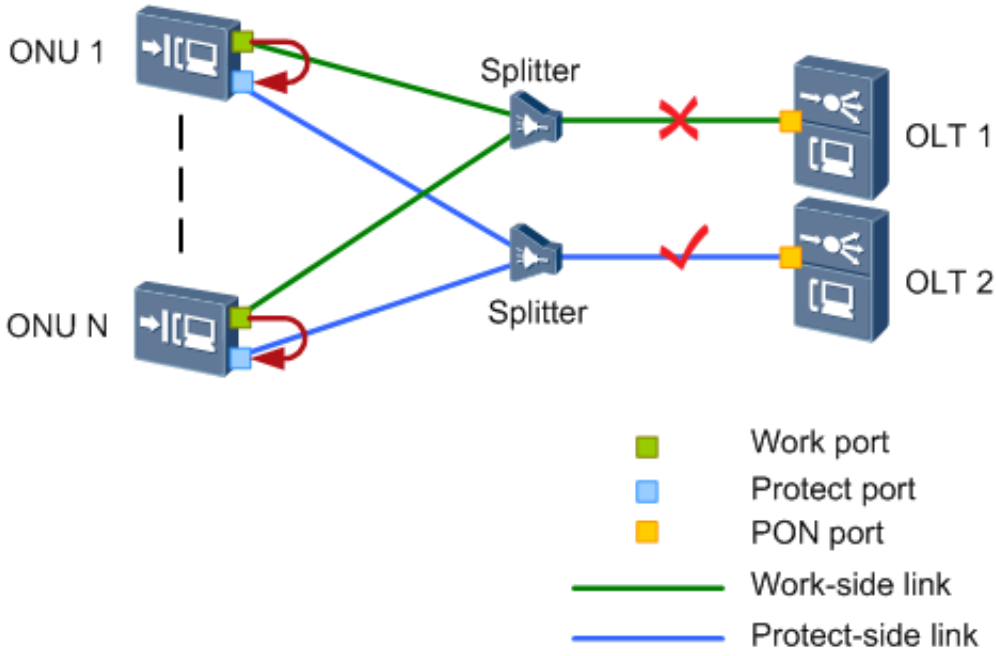


- If the ONU detects a fault on all the branch fibers connected to the working-side link, it automatically switches to the protection-side link and sends messages to OLT 2 to notify that protection switching has occurred. In addition, the ONU notifies OLT 2 of the switching cause.
- After switching, services on the ONU are transmitted to OLT 2 through the protection port (that is, service packets are transmitted by the protection-side link). The changes on the ONU are as follows:
  - The state of the working port changes to standby.
  - The state of the protection port changes to active.
- After protection switching, the ONU can automatically switch back to the working port. The OLT sends an automatic switchback message and the switchback time, called the WTR time, to the ONU. If the ONU learns that the working port, working-side links, and the uplink of OLT 1 are functioning properly and link A and link C stays normal during the WTR time, the ONU automatically switches to the working-side links when the WTR time expires.

Scenario 3: Backbone Fiber Becomes Faulty

Figure 2-38 shows the scenario in which the backbone fiber becomes faulty.

Figure 2-38 Backbone fiber becomes faulty

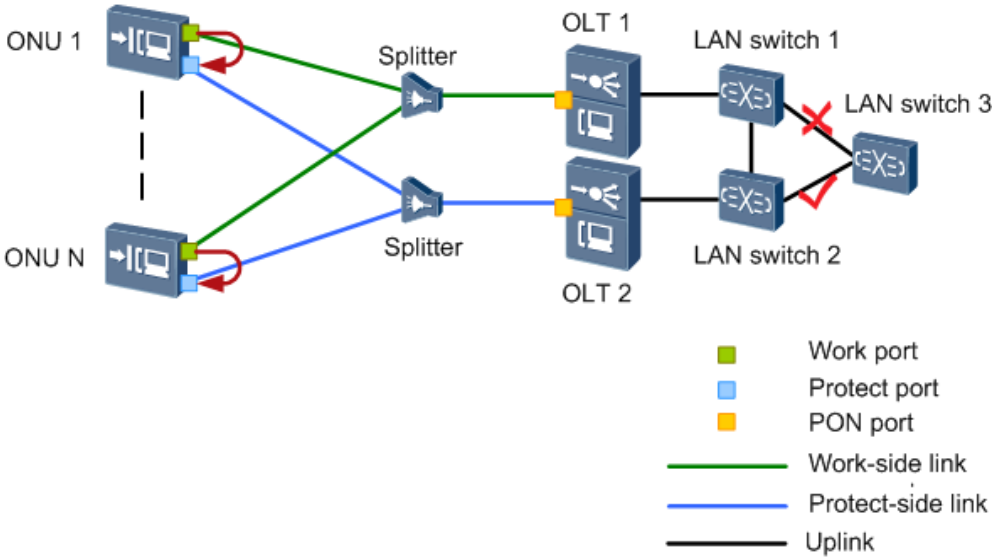


- If the ONU detects a fault on the backbone fiber, it automatically switches to the protection-side link and sends messages to OLT 2 to notify that protection switching has occurred. In addition, the ONU notifies OLT 2 of the switching cause.
- After switching, services on the ONU are transmitted to OLT 2 through the protection port (that is, service packets are transmitted by the protection-side link). The changes on the ONU are as follows:
  - The state of the working port changes to standby.
  - The state of the protection port changes to active.
- After protection switching, the ONU can automatically switch back to the working port. The OLT sends an automatic switchback message and the switchback time, called the WTR time, to the ONU. If the ONU learns that the working port, working-side links, and the uplink of OLT 1 are functioning properly and link A and link C stays normal during the WTR time, the ONU automatically switches to the working-side links when the WTR time expires.

Scenario 4: Associated Protection Switching Caused by a Fault on the OLT's Uplink

An OLT protection group is associated with the BFD or MEP session. Based on the association, when the upper-layer network connection (or IP layer link) of the OLT fails, the OLT instructs the ONU to trigger protection switching, which ensures service continuity. Figure 2-39 shows the associated protection switching caused by a fault on the OLT's uplink.

Figure 2-39 Associated protection switching caused by a fault on the OLT's uplink

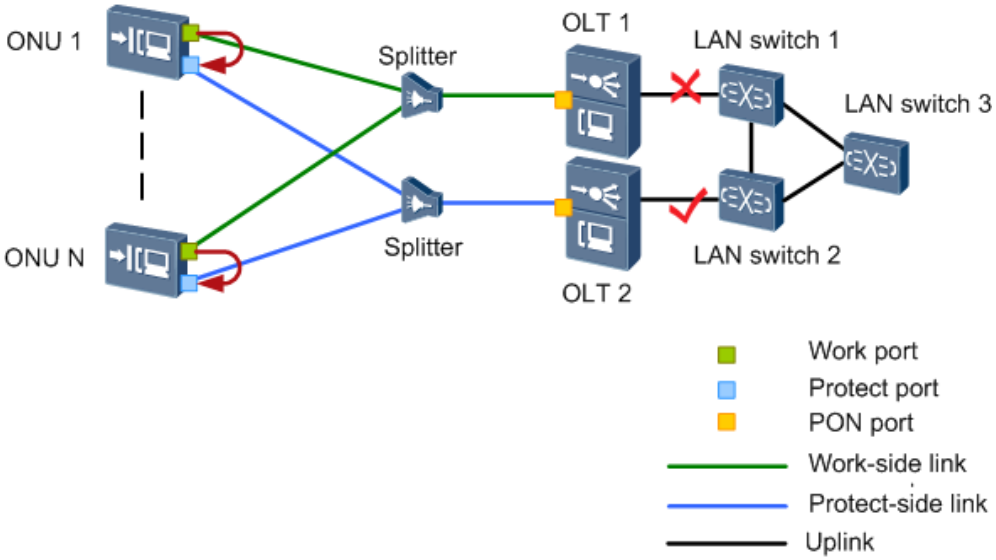


- On the OLT, the dual-homing protection group is associated with the BFD or MEP session. If the upstream route of the OLT fails, the BFD session status becomes **Down** or the MEP detection fails. In such a case, the OLT notifies the ONU of the uplink change.
- After the ONU receives the switching instruction from the OLT, it determines to trigger the switching and switches to the protection-side link. After the switching, the ONU notifies OLT 1 of the switching using the working-side link and notifies OLT 2 of the switching using the protection-side link.
- After switching, services on the ONU are transmitted to OLT 2 through the protection port (that is, service packets are transmitted by the protection-side link). The changes on the ONU are as follows:
  - The state of the working port changes to standby.
  - The state of the protection port changes to active.
- After protection switching, the ONU can automatically switch back to the working port. The OLT sends an automatic switchback message and the switchback time, called the WTR time, to the ONU. If the ONU learns that the working port, working-side links, and the uplink of OLT 1 are functioning properly and link A and link C stays normal during the WTR time, the ONU automatically switches to the working-side links when the WTR time expires.

Scenario 5: Associated Protection Switching Caused by a Fault on the OLT's Layer 2 Physical Link

An OLT protection group is associated with the uplink Ethernet port status. Based on the association, when the Layer 2 physical link of the OLT fails, the OLT instructs the ONU to trigger protection switching, which ensures normal service transmission. Figure 2-40 shows the associated protection switching caused by a fault on the OLT's Layer 2 physical link.

Figure 2-40 Associated protection switching caused by a fault on the OLT's Layer 2 physical link



- On the OLT, the protection group is associated with the uplink Ethernet port status. When the Ethernet port associated with the protection group becomes **Down**, the OLT notifies the ONU of the uplink change.
- After the ONU receives the switching instruction from the OLT, it determines to trigger the switching and switches to the protection-side link. After the switching, the ONU notifies OLT 1 of the switching using the working-side link and notifies OLT 2 of the switching using the protection-side link.
- After switching, services on the ONU are transmitted to OLT 2 through the protection port (that is, service packets are transmitted by the protection-side link). The changes on the ONU are as follows:
  - The state of the working port changes to standby.
  - The state of the protection port changes to active.
- After protection switching, the ONU can automatically switch back to the working port. The OLT sends an automatic switchback message and the switchback time, called the WTR time, to the ONU. If the ONU learns that the working port, working-side links, and the uplink of OLT 1 are functioning properly and link A and link C stays normal during the WTR time, the ONU automatically switches to the working-side links when the WTR time expires.

2.9 Remote Software Commissioning (GPON)

This section describes the implementation principles and configuration of remote software commissioning using GPON upstream transmission.

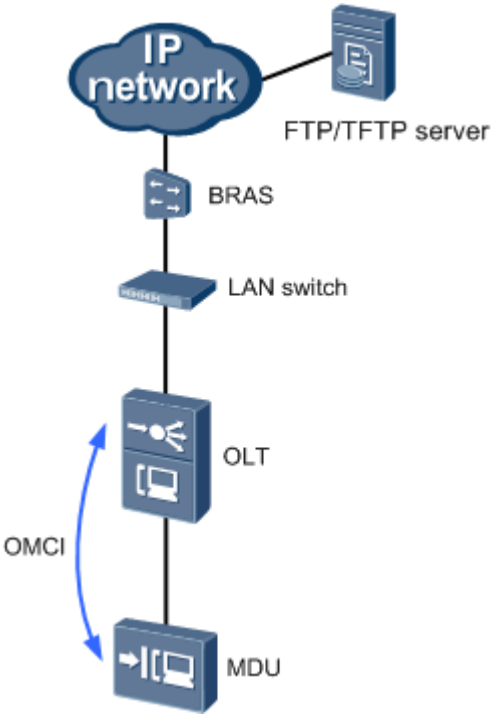
2.9.1 Introduction

During site deployment for a multi-dwelling unit (MDU) using GPON upstream transmission, the MDU can be functional only after it is installed and manually commissioned by commissioning engineers onsite. To remove the need for onsite MDU commissioning, the MDU supports remote software commissioning. After the MDU is powered on, it

automatically registers with the optical line terminal (OLT) and configures device data. This reduces site deployment costs.

Figure 2-41 shows the networking of remote software commissioning using GPON upstream transmission.

Figure 2-41 Networking of remote software commissioning using GPON upstream transmission



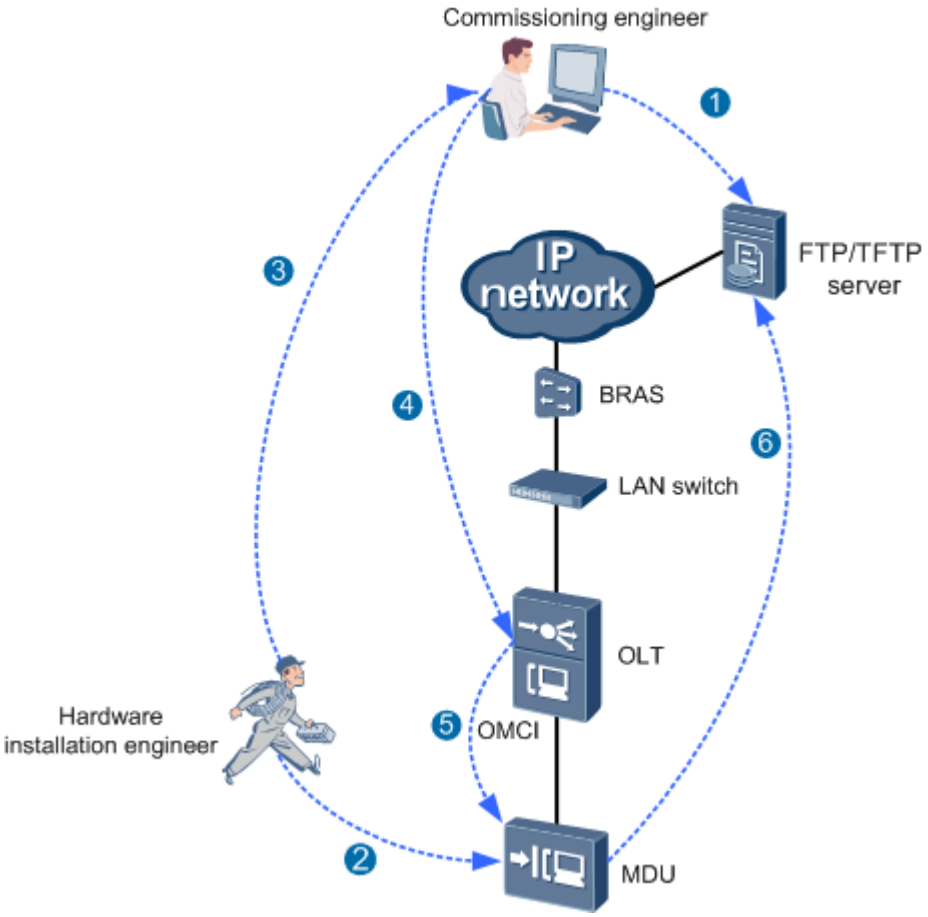
1. The OLT uses optical network terminal management and control interface (OMCI) to send the path where the automatic deployment policy file is stored to the MDU.
2. After being powered on, the MDU receives the path where the automatic deployment policy file is stored and starts automatic device deployment.
3. The MDU requests for the automatic deployment policy file from the FTP or TFTP server and implements automatic device configuration based on the automatic deployment policy specified in the file.

NOTE
Use FTP because it is more secure than TFTP.

2.9.2 Principles

Figure 2-42 shows the principles of remote software commissioning using GPON upstream transmission.

Figure 2-42 Principles of remote software commissioning using GPON upstream transmission



The process is as follows:

1. The commissioning engineer develops and uploads the automatic deployment policy file and configuration file to the FTP or TFTP server.

NOTE

The automatic deployment policy file must comply with the xxx.xml naming format. The file must contain the device type, control board, protocol for transferring the configuration file, IP address of the server, and configuration file name.

The configuration file name must be of string type.

One automatic deployment policy file applies to all MDUs in one site. An example automatic deployment policy file used in one site is as follows:

```
<config>
<deploy-version value="1"/>
<product name="MA5633">                    //Device type
<mainboard name="H822CCKRA"/>               //Control board
<mainboard name="H822CCKRB"/>
<mainboard name="H822CCKRC"/>
<load>
<transfer-protocol isSupport="1" protocol="ftp" username="user"
password="user123" port="" serveripaddr="10.10.10.10"/>
<init-load-script>
<common-script isSupport="1" value="script.txt"/>
```

```
</init-load-script>
</load>
</product>
</config>
```

- **protocol**: indicates the protocol for transferring the configuration file, which can be FTP, SFTP, or TFTP.
- **username** and **password**: indicate the user name and password, respectively, when the configuration file is transferred using FTP or SFTP.
- **port**: specifies a port. This parameter is required to configure only if the default port used by the transfer protocol must be changed.
- **serveripaddr**: indicates the server IP address.
- **value**: specifies a configuration file name. When the configuration file is transferred using FTP or SFTP, the configuration file name may contain the path where this file is stored.

2. The hardware installation engineer obtains the MDU from the warehouse and delivers it to the site. Then, the hardware installation engineer installs the MDU hardware, connects lines for the MDU, and powers on the MDU.
3. The hardware installation engineer records and reports the MAC address of the MDU and site information to the commissioning engineer.
4. The commissioning engineer adds the MDU to the OLT in offline mode and configures the IP address, service flows, and automatic deployment profile for this MDU.
5. After being powered on, the MDU receives the path where the automatic deployment policy file is stored and starts automatic device deployment.

NOTE

The automatic device deployment takes effect on the MDU only if the MDU starts from an empty database. If the MDU database is not empty, run the **erase flash data** command to clear the database, or run the **load data** command to load an empty database to the MDU.

6. The MDU requests for the automatic deployment policy file from the FTP or TFTP server and implements automatic device configuration based on the automatic deployment policy specified in the file.

2.9.3 Configuring Remote Software Commissioning (GPON)

The MDU supports remote software commissioning using GPON upstream transmission. After the MDU is powered on, it automatically registers with the OLT and configures device data.

Procedure

Run the **ont add** command add an MDU in offline mode.

Step 1 Run the **rn ipconfig** command to set the IP address of this MDU.

Step 2 Run the **service-port** command to create service flows.

Step 3 Run the **rn deploy-profile add** command to configure an automatic deployment policy profile.

NOTE

In remote software commissioning, the **terminal user authentication-mode AAA** *domain-name* command needs to be set at the last of the configuration file. Otherwise, this command configuration fails to be issued.

Step 4 Run the **rn deploy-config** command to bind the configured profile to the MDU.

Step 5 Run the **display rn deploy log** command to query automatic deployment results and failure causes if the deployment fails.

----End

Example

The following configurations are used as an example to configure the remote software commissioning feature:

1. Configure MDU 0 on port 0/14/0 as follows:
   - IP address of the MDU: 192.168.1.33
   - Subnet mask: 255.255.255.0
   - IP address of the gateway: 192.168.1.1
   - Management VLAN ID: 1
   - Priority: 3
2. Add automatic deployment policy profile 1 with automatic deployment policy file named **deploy-backup.xml**. The IP address of the file server is 10.10.10.10, the configuration file is transferred using FTP, and the user name and password are **user** and **user123**, respectively. Bind the automatic deployment policy profile to port 0/14/0.

```
huawei(config)#interface gpon 0/14
huawei(config-if-gpon-0/14)#port 0 ont-auto-find enable
huawei(config-if-gpon-0/14)#ont add 0 sn-auth 485754437B6F5130 snmp ont-lineprofile-id
1
huawei(config-if-gpon-0/14)#quit
huawei(config)#rn ipconfig 0/14/0 0 ip-address 192.168.1.33 mask 255.255.255.0 gateway
192.168.1.1 vlan 1 priority 3
huawei(config)#service-port vlan 1 gpon 0/14/0 ont 0 gemport 0 multi-service user-vlan
1
huawei(config)#rn deploy-profile add profile-id 1 filename deploy-backup.xml ip
10.10.10.10 ftp user
huawei(config)#rn deploy-config 0/14/0 0 profile-id 1
```

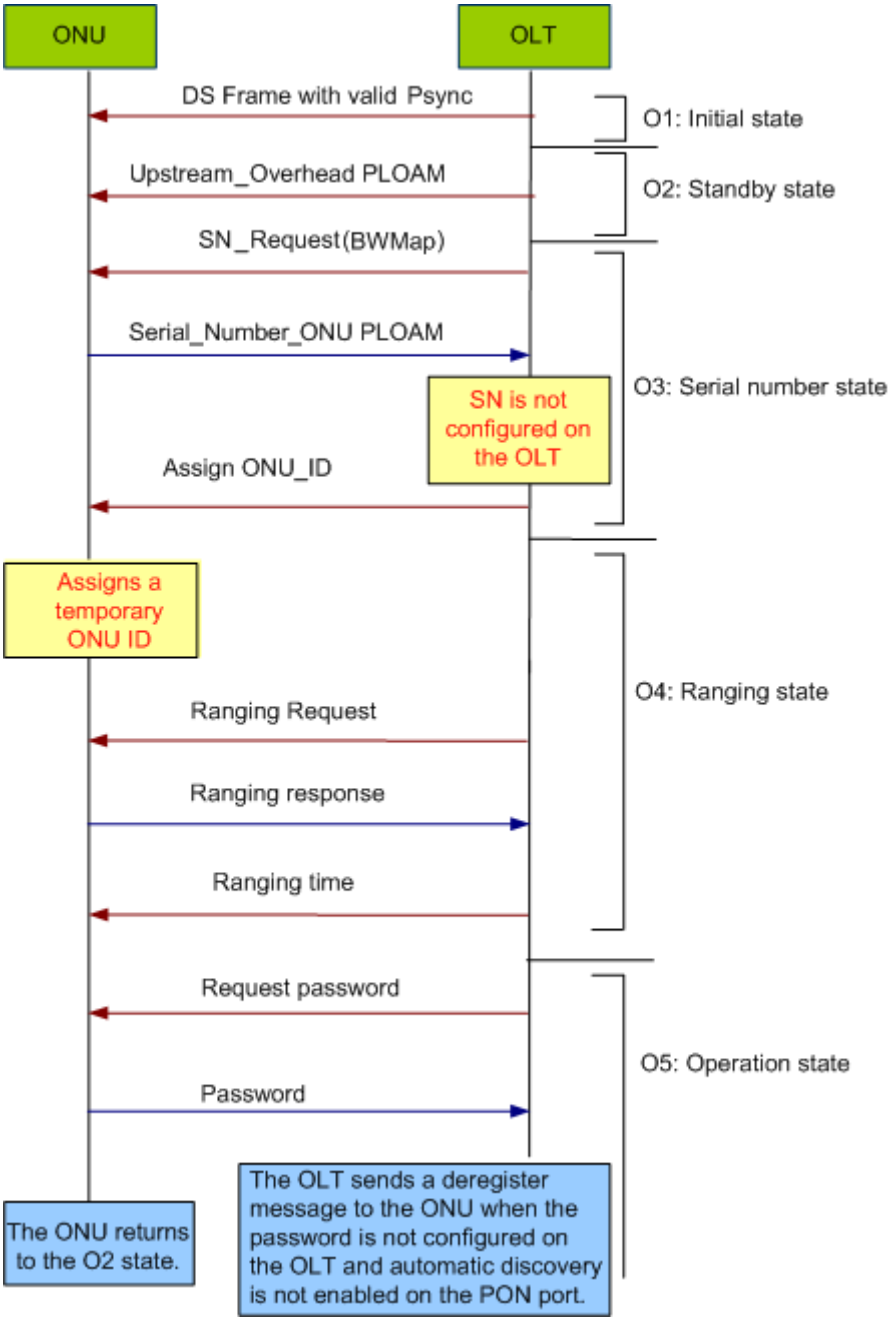
2.10 GPON Terminal Authentication and Management

GPON terminal authentication is a mechanism in which an OLT authenticates an ONU according to the authentication information reported by the ONU and in this way denies access to unauthorized ONUs. In the GPON system, only authenticated ONUs can access the system. After the ONU passes authentication and goes online, data can be transmitted between ONUs and the OLT.

2.10.1 GPON Terminal Authentication (ONU Is Not Preconfigured)

Figure 2-43 shows the authentication process of an ONU that is not preconfigured.

Figure 2-43 Authentication process of an ONU that is not preconfigured



1. The OLT sends an serial number (SN) request to the ONU.
2. The ONU responds to the SN request message sent from the OLT.
3. Upon receiving the SN response from the ONU, the OLT assigns a temporary ONU ID to the ONU.
4. After the ONU enters the operation state, the OLT sends a password request message to the ONU. The ONU then responds with a password. The password is not configured on the OLT.

- If the automatic discovery function is not enabled on the PON port to which the ONU is connected, the OLT sends a deregister message to the ONU. Upon receiving this message, the ONU sends a register request message to the OLT.
- If the automatic discovery function is enabled on the PON port to which the ONU is connected, the port reports an alarm to the command line interface (CLI) or network management system (NMS), indicating that the ONU is automatically discovered. The ONU can go online only after being confirmed.

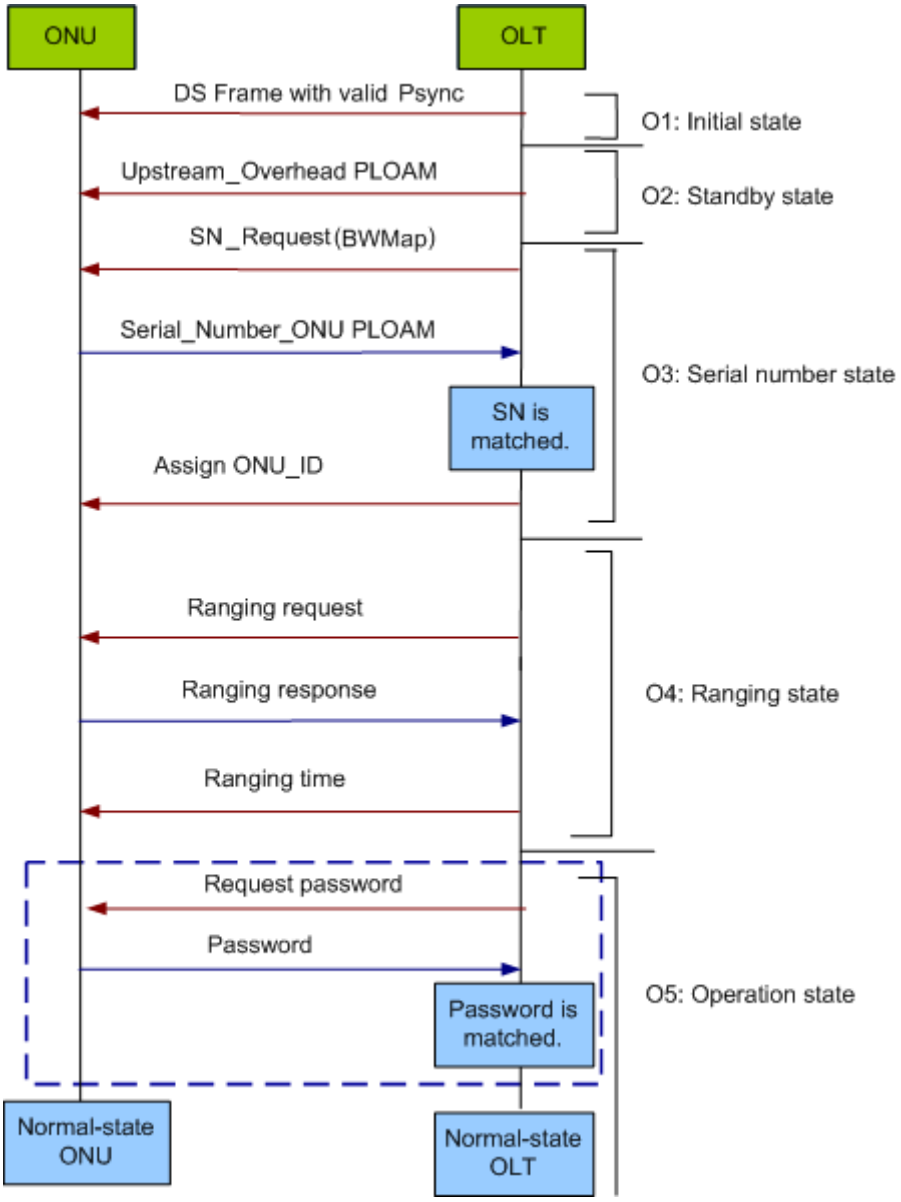
2.10.2 GPON Terminal Authentication (ONU Has Been Pre-configured)

A pre-configured ONU can be authenticated in three modes: SN, SN+password, and password.

SN/SN+Password Authentication

In SN authentication, the OLT matches only the ONU SN. In SN+password authentication, the OLT matches both the ONU SN and password. Figure 2-44 shows the authentication flow.

Figure 2-44 SN/SN+password authentication flow



NOTE

If an ONU is authenticated in SN mode, no password is required in the authentication process.

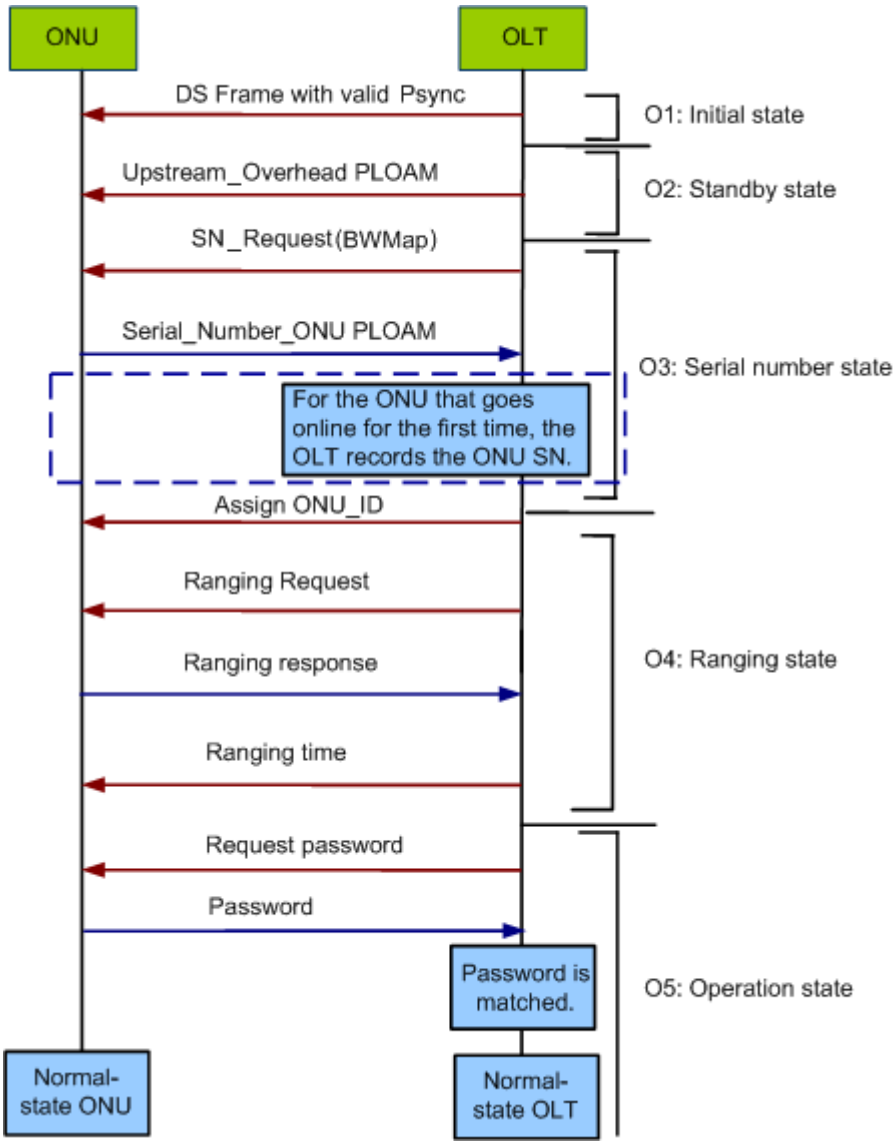
- After receiving an SN response message from an ONU, the OLT checks whether another ONU with the same SN is online. If yes, the OLT reports an SN conflict alarm to the CLI or NMS. If no, the OLT directly assigns a user-defined ONU ID to the ONU.
- After the ONU enters the operation state,
  - For the ONU that is authenticated in SN mode, the OLT does not send a password request message to this ONU. Instead, the OLT automatically configures a GEM port that has the same ID as the ONU ID for the ONU for carrying OMCI messages, and allows the ONU to go online. In addition, the OLT reports an ONU online alarm to the CLI or NMS.

- For the ONU that is authenticated in SN+password mode, the OLT sends a password request to the ONU, and compares the password reported by the ONU with the local password. If the two passwords are the same, the OLT directly configures a GEM port for the ONU to carry OMCI messages, and allows the ONU to go online. In addition, the OLT reports an ONU online alarm to the CLI or NMS. If the two passwords are not the same, the OLT reports a password error alarm to the CLI or NMS. The OLT does not report an ONU automatic discovery message even if the ONU automatic discovery function is enabled on the PON port. Instead, the OLT sends the Deactivate_ONU-ID PLOAM message to deregister the ONU.

Password Authentication

An ONU that uses password authentication is added to a PON port on an OLT in advance, and then this ONU is connected to the PON port. In password authentication, if finding that the SN or password of the ONU to be authenticated conflicts with that of an online ONU, the OLT deregisters the ONU to be authenticated. This does not affect the online ONU. Password authentication is available in two modes: once-on and always-on.

Figure 2-45 Initial ONU authentication in once-on mode



NOTE

During the authentication in always-on mode, the OLT does not need to record the SN of the ONU that goes online for the first time.

Once-on Application Scenarios

A carrier allocates a password to a user and requires the user to go online within a specified time. After going online, the user cannot change the ONU. To change the ONU, the user must notify the carrier. In once-on mode, the aging time is configurable. After the aging time is set, the ONU must register with the OLT and go online within the preset aging time. Otherwise, the ONU is not allowed to register with the OLT or go online. Once the ONU is authenticated, its SN cannot be changed.

In once-on mode,

- Only the initial authentication of an ONU is performed by password, as shown in Figure 2-45.

- In subsequent authentications, the ONU can be authenticated by SN or SN+password according to the CLI configuration, as shown in Figure 2-44.

NOTE

In once-on mode, before the ONU registration times out or before the ONU successfully registers with the OLT for the first time, the ONU discovery status is ON. Only the ONU whose discovery status is ON is allowed to register with the OLT and go online. After the ONU registration times out or after the ONU successfully registers with the OLT for the first time, the OLT sets the ONU discovery status to OFF.

- The ONU whose registration times out is not allowed to register with the OLT or go online. The registration timeout flag of the ONU needs to be reset at the central office (CO), and then the ONU can go online.
- An ONU that successfully registers for the first time is allowed to register and go online again.

Always-on Application Scenarios

The always-on mode applies to the following scenario: A carrier allocates a password to a user, and the user can use different ONUs with this password and different SNs. The user can change the ONU without informing the carrier. In always-on mode, there is no restriction on the time when the user goes online.

- An ONU is authenticated by password when it goes online for the first time. After the ONU passes the password authentication and goes online successfully, the OLT generates an SN+password entry according to the SN and password of the ONU. Figure 2-45 shows the authentication process.
- The following scenarios are involved if it is not the first time that an ONU goes online:
  - If the SN and password of the ONU are the same as the SN and password of the ONU that successfully goes online for the first time, the ONU is authenticated by SN+password. Figure 2-44 shows the authentication process.
  - If the user replaces the ONU with an ONU that has the same password but a different SN, the new ONU is authenticated by password. After this ONU passes authentication and goes online successfully, the original SN+password entry is updated. Figure 2-45 shows the authentication process.

2.10.3 GPON Terminal Management

The ONUs in a GPON system are managed using physical layer OAM (PLOAM) messages and OMCI messages.

PLOAM, defined in ITU-T Recommendation G.984.3, is used for exchanging management and maintenance messages, such as DBA and DBRu messages, between the GPON physical layer and TC layer.

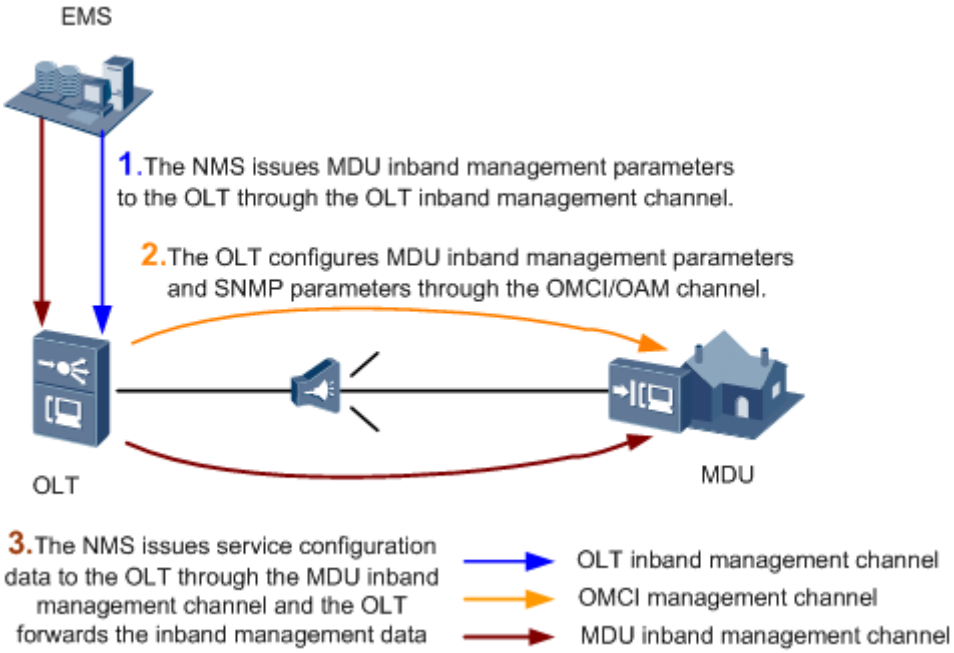
GPON ONUs, including MDUs and ONTs, are managed using OMCI messages. The ONUs are plug and play and support offline deployment and automatic service provisioning. For details about OMCI management functions, see 2.6.4 OMCI.

- OMCI messages are used for maintaining and managing service hierarchies, such as discovering device hardware capabilities and configuring alarm maintenance and service capabilities.
- OMCI enables ONUs to support offline configuration so that the ONUs do not need to store configuration data locally, facilitating service provisioning.

MDU Management

Figure 2-46 shows the process of configuring a management channel for an MDU.

Figure 2-46 Process of configuring a management channel for an MDU



1. The NMS issues MDU inband management parameters to the OLT through the OLT inband management channel.
2. The OLT configures the MDU inband management parameters and Simple Network Management Protocol (SNMP) parameters through the OMCI or OAM channel to set up the MDU inband management channel.
3. The NMS issues service configuration data through the MDU inband management channel. After the MDU inband management channel is set up, the NMS configures and manages the MDU through the SNMP channel. In such a manner, the OLT only needs to forward the MDU inband management data.

ONT Management

GPON terminals are managed using one of these protocols: optical network terminal management and control interface (OMCI), Extensible Markup Language (XML), or Technical Report 069 (TR069).

- The optical network terminal management and control interface (OMCI) protocol is defined by ITU-T G.984.4, which applies to managing optical network terminals (ONTs) in a GPON system. Huawei ONTs comply with OMCI. OMCI messages are transmitted between an optical line terminal (OLT) and an ONT over a dedicated permanent virtual channel (PVC) in asynchronous transfer mode (ATM) or a GPON encapsulation mode (GEM) port. The OMCI protocol manages and provides O&M for the ONT.
- Extensible Markup Language (XML) is a text format used for message interaction between devices. The iManager U2000 Unified Network Management System (U2000) uses XML to manage ONTs in a Huawei FTTx system. XML is also a management mode extended from OAM because not all voice and Layer 3 gateway services are defined in the OAM.
- Technical Report 069 (TR069) is a network management protocol defined by the DSL Forum. The full name of TR069 is CPE WAN Management Protocol (CWMP). CPE is the acronym for customer premises equipment and WAN is the acronym for wide area

network. TR069 defines a new network management structure consisting of management models, interaction interfaces, and basic management parameters. In the network management structure, the management server functions as an Auto-Configuration Server (ACS) and is responsible for managing the CPE. The ACS and CPE use Hypertext Transfer Protocol (HTTP) to communicate with each other. The ACS serves as an HTTP server and the CPE serves as a client. Management operations are implemented using XML-based remote procedure call (RPC).

Optical network terminals (ONTs) are classified into three types: bridge type, bridge+voice type, and gateway type.

- A bridge-type ONT provides Layer 2 data and multicast services.
- A bridge+voice-type ONT provides Layer 2 data, Layer 2 multicast services, and voice over IP (VoIP) services.
- A gateway-type ONT provides Layer 3 data, Layer 3 multicast services, and VoIP services.

Each different type of terminal management protocol has a unique service management scope. Based on terminal types, provides three GPON terminal management solutions: OMCI, OMCI+XML, and OMCI+TR069. The advantages and disadvantages of each solution as well as the recommended solution for each type of ONT are listed at the end of this chapter.

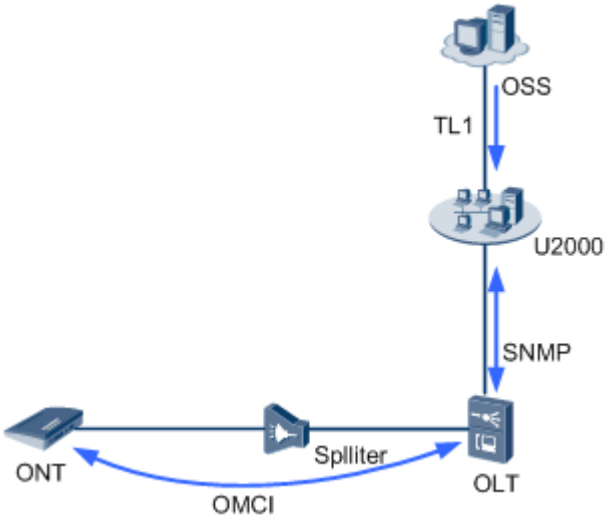
- The OMCI protocol manages Layer 2 services, voice services and the PON link layer. This protocol cannot manage Layer 3 services.
- The XML protocol manages Layer 3 services and voice services. Using OMCI+XML enables you to manage Layer 2, voice, and Layer 3 services.
- The TR069 protocol manages Layer 3 services and voice services, and identifies remote faults. When this protocol is used, OMCI is still used to manage Layer 2 services and the PON link layer.

OMCI

A standard optical network terminal management and control interface (OMCI) solution enables you to manage optical network terminals (ONTs) supplied by different vendors in diverse types of scenarios. An optical line terminal (OLT) and an ONT are closely coupled with each other. If a new service requirement is not defined in the OMCI, a new OMCI entity must be defined. An OMCI solution enables you to manage Layer 2 features and voice services. The OLT communicates with the ONT in OMCI mode.

Figure 2-47 shows the general principles of the OMCI solution for U2000+OLT+ONT deployment scenarios.

Figure 2-47 General principles of the OMCI solution



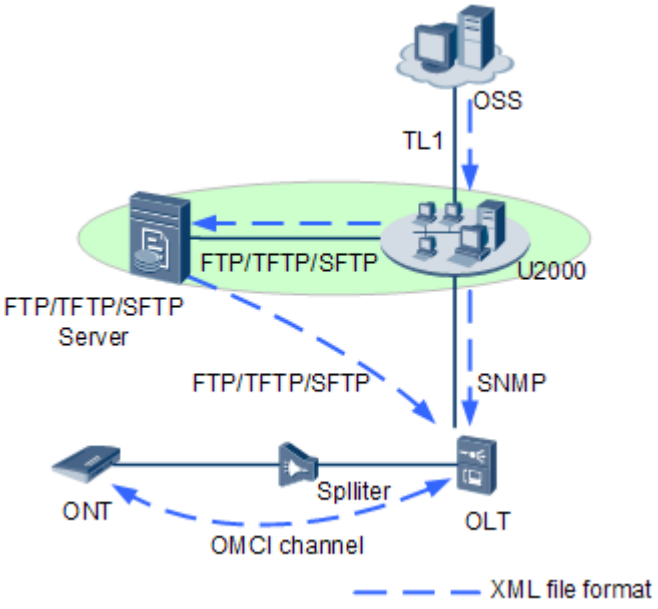
1. The Operations Support System (OSS) issues service configuration parameters to the iManager U2000 Unified Network Management System (U2000) using the TL1 northbound interface (NBI).
2. The U2000 uses Simple Network Management Protocol (SNMP) to manage the OLT.
3. The OLT issues service configuration parameters to the ONT through an OMCI channel.

XML+OMCI

To overcome the limitations of the OMCI solution, Huawei provides a solution that combines the XML protocol with the OMCI protocol. In the XML+OMCI solution, the U2000 uses XML files transmitted over an IP channel to communicate with the OLT, and the OLT uses XML files transmitted over an OMCI channel to communicate with the ONT. The OMCI protocol manages Layer 2 services and the XML protocol manages Layer 3 and voice services.

Figure 2-48 shows the general principles of the XML+OMCI solution for U2000+OLT+ONT deployment scenarios.

Figure 2-48 General principles of the XML+OMCI solution



As part of the general principles, the U2000 uploads XML files to a File Transfer Protocol (FTP)/Trivial File Transfer Protocol (TFTP)/Secure File Transfer Protocol (SFTP) server. Then the OLT obtains the XML files from the FTP/TFTP/SFTP server and transparently transmits the files to the ONT through the OMCI channel.

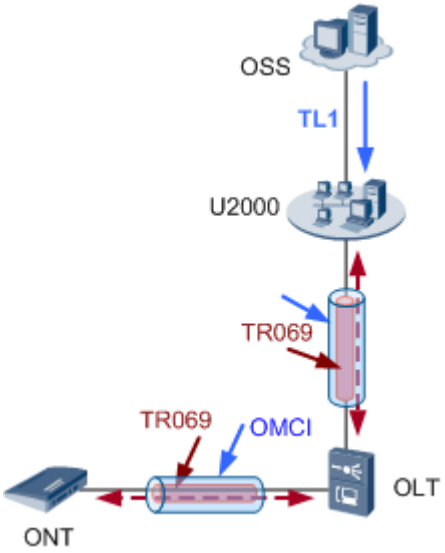
NOTE
SFTP loading is recommended to load a XML files for an ONT.

1. The OSS issues service configuration parameters to the U2000 using the TL1 NBI.
2. The U2000 converts service information to XML files and uploads the files to the FTP/TFTP/SFTP server.
3. The U2000 issues ONT configuration update commands to the OLT and asks the OLT to download the files.
4. The OLT obtains the XML files from the FTP/TFTP/SFTP server.
5. The OLT issues the XML files to the ONT through the OMCI channel.
6. The ONT returns execution results to the OLT using the OMCI entity.
7. The OLT reports the results to the U2000 in traps.

The XML+OMCI solution meets all requirements for configuring the ONT but configuration files are transmitted in unidirectional mode. Due to this limitation, the configuration files only implement service configurations and status performance management, but cannot provide operation and maintenance (O&M) functions such as query of ONT status and configuration, and test and diagnose functions. To overcome XML+OMCI limitations, Huawei provides TR069 over OMCI. As a supplement to XML+OMCI, TR069 over OMCI is used for remote O&M and fault identification. The U2000 can use TR069 to remotely maintain the ONT without a dedicated TR069 server.

Figure 2-49 shows the general principles of the TR069 over OMCI solution for U2000+OLT+ONT deployment scenarios.

Figure 2-49 General principles of the TR069 over OMCI solution



The solution manages configuration, performance, faults, and status of IP-based services by applying the associated methods described in the TR069 solution to the OMCI solution. The OLT and ONT transparently transmit data between each other.

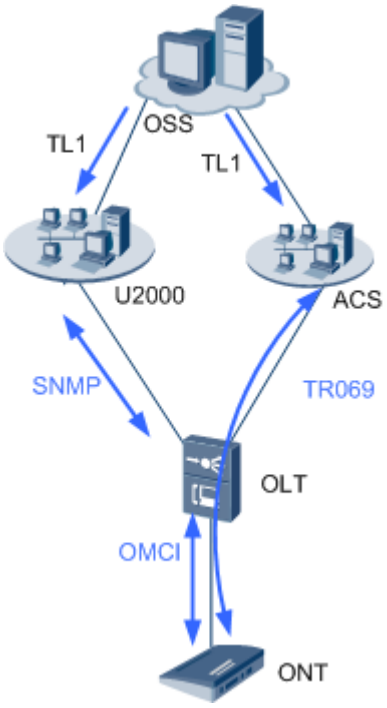
1. The U2000 manages and maintains the ONT, and queries the ONT status. The U2000 encapsulates management, maintenance, and query data to character strings or binary code streams in a specific format and sends them to the OLT through a management information base (MIB) interface.
2. The OLT transparently transmits the character strings or binary code streams to the ONT using an extended OMCI entity.
3. The ONT returns execution results to the OLT using the OMCI entity.
4. The OLT reports the results to the U2000 in traps.

OMCI+TR069

This solution allows an Auto-Configuration Server (ACS) to manage all the terminals on the network, locate faults, provide services, and collect performance statistics. Based on SNMP and TR069, this solution allows the ACS to manage home terminals in a unified manner, reducing O&M costs. TR069 automatically implements ONT configuration, dynamically provisions services, remotely locates faults, and rapidly collects terminal statistics.

Figure 2-50 shows the general principles of the OMCI+TR069 solution.

Figure 2-50 General principles of the OMCI+TR069 solution



This solution allows the U2000 to manage the OLT using SNMP, manage voice and Layer 3 services using TR069, and manage PON link layer using OMCI.

1. The OSS issues service configuration parameters to the U2000 using the TL1 NBI.
2. The U2000 manages the OLT using SNMP.
3. The OLT issues PON link layer configuration to the ONT using OMCI.
4. The ONT returns execution results to the OLT. Then the IP channel is set up.
5. The ONT registers with the ACS.
6. The ACS encapsulates user information in a TR069-compliant format and sends it to the ONT through the IP channel. The user information includes operations, maintenance items, and queries performed by a user. The IP channel is bidirectional.

Advantages and Disadvantages of the Terminal Management Solutions

Bridge type, bridge+voice type, and gateway type ONTs provide different types of services. Therefore, different solutions are used to manage these ONTs. Table 2-8 lists the advantages and disadvantages of each solution. Table 2-9 lists the recommended solution for each type of ONT.

Table 2-8 Advantages and disadvantages of each solution

Terminal Management Solution	Advantage	Disadvantage
OMCI	• A unified interface is used for ONT service management.	• The OLT and ONT are closely coupled with on each other. New services on the ONT require the

Terminal Management Solution	Advantage	Disadvantage
	• The OLT and ONT communicate with each other using OMCI-associated standards. • The ONT does not require a management IP address.	OLT's support, adding to the difficulty in deploying new services. • The OMCI standard is not fully developed. If a new service requirement is not defined in the OMCI, a new OMCI entity must be defined.
OMCI+XML	• The ONT does not require a management IP address. • The OLT and ONT are not closely coupled with each other to certain extent. • A unified management server is used for swift service deployment.	• This is a Huawei's proprietary solution and cannot interact with devices from other vendors. • Voice and Layer 3 services cannot be configured using a command on the OLT.
OMCI+TR069	An OLT version and an ONT version are not bound to each other. In other words, an OLT upgrade does not require an ONT upgrade; the opposite is also true. TR069 provides an enhanced definition and deployment scenario for the IP-based customer premises equipment (CPE) service management model. Therefore, ONT vendors can easily deploy new gateway and voice services.	• TR069 is based on the IP protocol and requires an extra IP management network. • Different interfaces are used to manage the ONT. The network management system (NMS) manages the link layer and the ACS manages IP-based services.

Table 2-9 Recommended solutions for each type of ONT

Terminal Type	Optional Solution	Recommended Solution
Bridge type	OMCI	OMCI
Bridge+voice type	OMCI+XML or OMCI	OMCI+XML (NMS provisions services) OMCI (OLT is connected to the third-party ONT)
Gateway type	OMCI+XML or OMCI+TR069	OMCI+TR069

2.11 Continuous-Mode ONU Detection

Overview

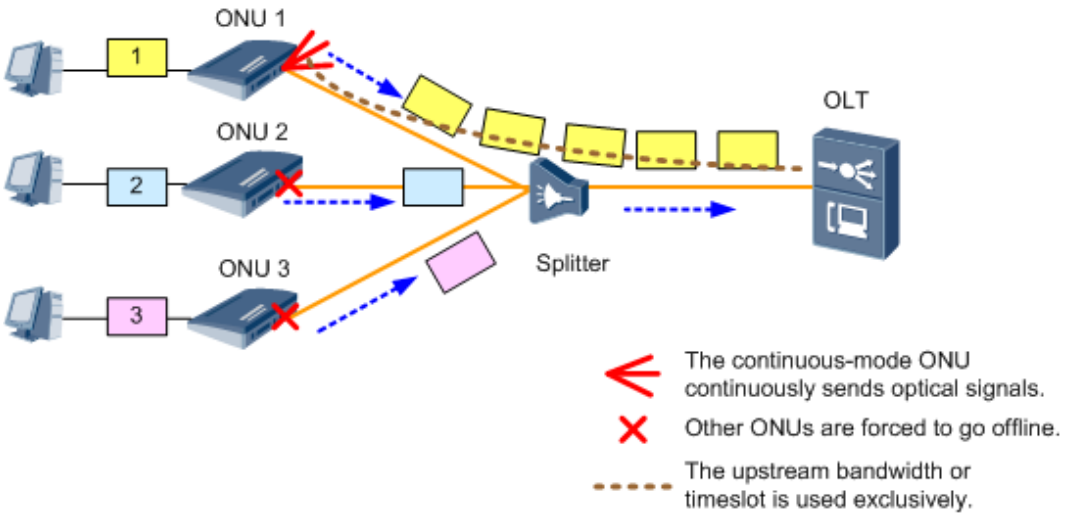
GPON networks use the P2MP network architecture. They use time division multiple access (TDMA) in the upstream direction. ONUs must send optical signals upstream at the timeslots allocated by the OLT to prevent data conflict.

The ONUs sending optical signals upstream not at the timeslots allocated by the OLT are continuous-mode ONUs, also called rogue ONUs. A continuous-mode ONU continuously sends optical signals.

A continuous-mode ONU adversely affects the system as follows:

- If this ONU has been online, some or all ONUs connected to the same PON port go offline or frequently go offline and online.
- If this ONU has not been configured, other ONUs that have not been configured and connected to the same PON port will not be discovered by the OLT.

Figure 2-51 Continuous-mode ONU



Continuous-Mode ONU Detection

Continuous-mode ONU detection, also called rogue ONU detection, is used for detecting continuous-mode ONUs in the system and isolating them, ensuring proper system running. A continuous-mode ONU detection process involves three stages, checking, detection, and isolation.

The three stages are as follows:

- Checking: The OLT periodically checks whether a continuous-mode ONU connects to a PON port. This checking cannot locate the continuous-mode ONU.

The OLT opens an empty gate in the upstream direction to detect ONU optical signals in the upstream direction. If the OLT receives optical signals, it then goes to the detection stage to locate the ONU.

- Detection: The OLT locates the continuous-mode ONU.

  The OLT issues a command to ONUs to instruct the optical modules of the ONUs to send optical signals upstream and checks whether optical signals can be received in the upstream direction. If other ONUs go offline after an ONU sends optical signals, this ONU is a continuous-mode ONU. In a detection process, the OLT checks all ONUs connected to a PON port for detecting all continuous-mode ONUs.

- Isolation: The OLT issues a command to power off the continuous-mode ONU, preventing this ONU from adversely affecting other ONUs connected to the same PON port.

  After an ONU is powered off by the OLT, the ONU cannot send optical signals upstream even after being reset or power recycled. This ONU can send optical signals upstream only after the OLT cancels the isolation.

NOTE
The OLT checks continuous-mode ONUs but does not detect or isolate them by default.

Handling a Continuous-Mode ONU

1. If an ONT goes online and other ONTs connected to the same PON port go offline or go online and offline frequently, or the 0x2e314021 There are illegal incursionary rogue ONTs under the port alarm is reported to the OLT, a rogue ONT may exist in the system. In this case, locate the rogue ONT according to the following steps.

NOTE
You can also run the **display port state** command to query whether a rogue ONT exists under a PON port.

2. Run the **anti-rogueont manual-detect** command to detect, locate, and isolate a continuous-mode rogue ONT manually. Then, check whether the system generates the The ONT is rogue ONT or There are illegal incursionary rogue ONTs under the port alarm.

NOTE
When you detect a rogue ONT, if a type B protection group is configured on the port that is connected to the ONT to be detected, you need to run the **force-switch** command to forcibly switch the protection group and then detect the rogue ONT to ensure that protection group switching does not occur during rogue ONT detection. You can forcibly switch services to the work side for rogue ONT detection if you are not sure which backbone fiber functions properly. If the rogue ONT is not detected, forcibly switch services to the protect side for rogue ONT detection. Then, run the **undo force-switch** command to cancel forced protection group switching.

   - If the The ONT is rogue ONT or There are illegal incursionary rogue ONTs under the port alarm is generated, a continuous-mode rogue ONT may exist. In this case, go to 3.
   - If the The ONT is rogue ONT or There are illegal incursionary rogue ONTs under the port alarm is not generated, an irregular-mode rogue ONT may exist. In this case, go to 4.

3. Handle the ONT according to the generated alarm.
   - If the The ONT is rogue ONT alarm is generated, replace the ONT. Then, go to 7.
   - If the There are illegal incursionary rogue ONTs under the port alarm is generated, go to 4.

NOTE

If the There are illegal incursionary rogue ONTs under the port alarm is generated, a continuous-mode ONT may exist and this ONT does not support Huawei-defined extended PLOAM messages or optical signal transmission of the ONT optical module cannot be controlled.

4. Run the **ont reset** command or the **ont deactivate** command to reset or deactivate ONTs under the PON port one by one. Then, check whether other ONTs that encounter the fault (going offline or going online and offline repeatedly) can go online.
   - If other ONTs that encounter the fault can go online, the ONT is a rogue ONT. Go to 7.
   - If other ONTs that encounter the fault cannot go online, the ONT optical module may be damaged so that the rogue ONT fails to be reset or deactivated by running the command. In this case, go to 5.
5. Locate a rogue ONT manually: On the optical splitter, remove upstream optical fibers of the ONTs one by one and check whether other ONTs that encounter the fault (going offline or going online and offline repeatedly) can go online.
   - If other ONTs that encounter the fault can go online, the ONT is a rogue ONT. Then, go to 7.
   - If other ONTs that encounter the fault cannot go online, the optical module may be damaged so that the rogue ONT fails to be reset or deactivated. In this case, go to 6.
6. Contact Huawei technical support.
7. The fault is rectified.

Limitations and Restrictions

- The OLT checks and analyzes the abnormality in the sending of upstream optical signal over a PON line, and identifies and isolates rogue ONUs of only non-malicious users. This feature does not apply to the intentionally sabotaged ONU or sub-standard ONU.
- A continuous-mode ONU (rogue ONU) is required to parse and respond to downstream PLOAM messages.
- When detecting a continuous-mode ONU, the OLT can quickly locate the continuous-mode ONU only if this ONU supports Huawei proprietary messages in the upstream direction.

2.12 Introduction to eOTDR

Context

Carriers face the following issues in different phases of PON O&M due to the lack of effective methods:

- The efficiency of PON service provisioning is low because no method is available to ensure smooth service provisioning.
- The fault report rate is high due to service faults resulting from deterioration and high attenuation of optical fibers because no method is available to monitor optical fiber status after the service is provisioned.
- Troubleshooting is time-consuming and customer satisfaction rate is low because no tool is available to locate an ODN network fault and therefore, the O&M personnel can only locate the fault segment by segment.

The OTDR test function provides the following benefits to carriers:

- Before a service is provisioned, the OTDR test checks optical fiber attenuation. Based on the test results, carriers can determine whether the service can be provisioned on the optical fiber.
- During network running, the OTDR test periodically checks optical fiber status to identify latent optical fiber faults, which can be rectified before they affect user services. This feature reduces the fault report rate.
- After a user reports a fault, the OTDR test identifies an optical fiber fault, such as optical fiber cut, or deterioration, and provides the location of the fault, such as indoor, outdoor, inside a building, or outside a building. Based on the test results, the O&M personnel can correctly send a dispatch for troubleshooting, which reduces O&M costs.

OTDR Principles

TDR is the acronym of optical time domain reflectometer. An OTDR is an optical and electrical integrated meter based on the backscattering generated as a result of Rayleigh scattering and Fresnel reflection during the light transmission over optical fibers. It is widely used in the maintenance and engineering of optical cables. Engineers can use the OTDR to test the length, transmission attenuation, and connector attenuation of optical fibers and locate optical fiber faults.

NOTE
For more information about OTDR principles, see the N2510 documentation.

eOTDR Principles

An eOTDR is an embedded OTDR, with OTDR functions integrated into a data communication optical module. An eOTDR enables an optical module used in a PON network to integrate the functions of a data transmitter, data receiver, OTDR transmitter, and OTDR receiver. OTDRs use a separate test wavelength. Figure 2-52 shows eOTDR applications.

Figure 2-52 eOTDR applications

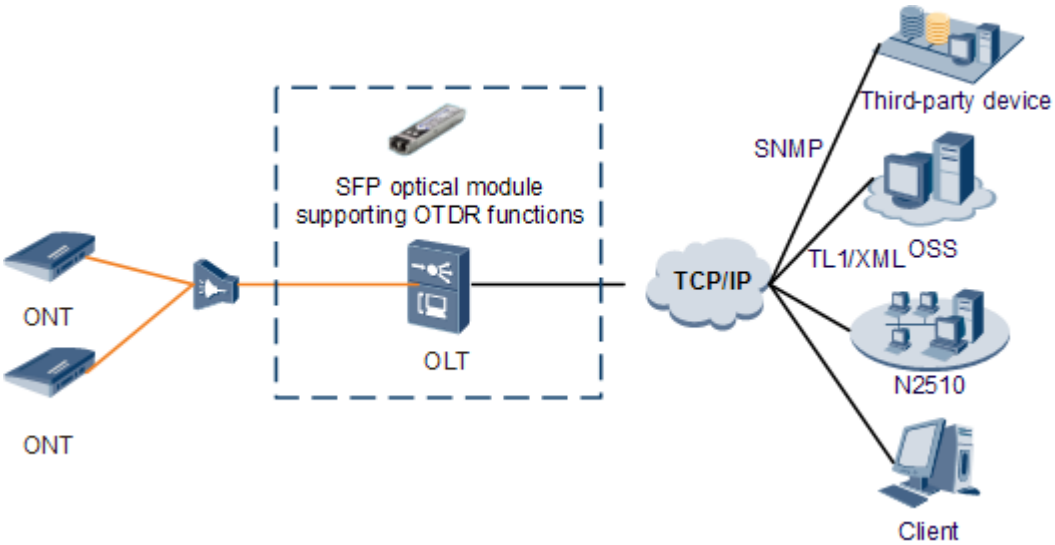
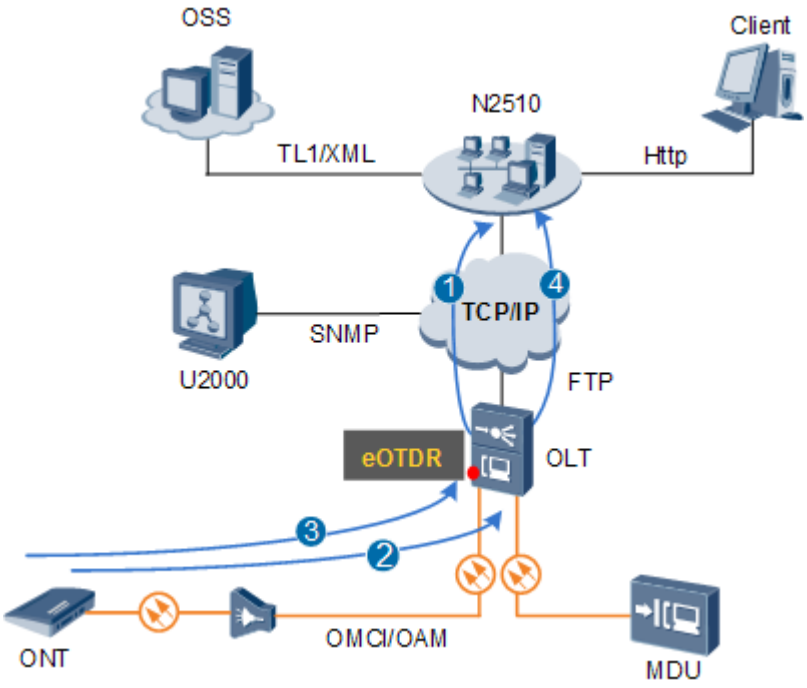


Figure 2-53 shows eOTDR test procedure.

Figure 2-53 eOTDR test procedure



1. The N2510 sends test commands and parameters to the eOTDR of the OLT using SNMP.
2. The eOTDR starts the test.
3. The OLT collects test results.
4. The OLT sends the test results to the N2510. Then, the N2510 analyzes the data.

eOTDR Highlights

- Precise and efficient fault demarcation, reducing possibility of wrong dispatching of orders
  - Demarcates fiber faults of CO, ODN (feeder fiber, distribution fiber, and optical distribution point), and ONTs.
  - Demarcates fiber faults inside/outside users' houses and buildings.
- Precise fault location, improving fault rectification efficiency
  - Precisely differentiates between fault types (excessively small fiber bending radius, fiber cut, and air gap).
  - Graphically displays information in GIS.
- Proactive OAM, reducing customer complaints
  - Monitors fiber line performance.
  - Diagnoses lines based on alarms.
  - Analyzes network quality.

eOTDR Specifications

Table 2-10 shows eOTDR hardware specifications.

Table 2-10 eOTDR hardware specifications

Item	Specifications
Test wavelength	1490 nm
Test range	20 km
Test pulse width	12.5 ns, 25 ns, 50 ns, 100 ns, 200 ns, 400 ns, or 800 ns
Dynamic range	7 dB@100 ns or 5 dB@25 ns
Reflection event dead zone	5 m@25 ns
Attenuation event dead zone	50 m@25 ns
Start dead zone	80 m@25 ns

Table 2-11 eOTDR detection capability (1:8 split ratio)

Item	Detection Capability
Reflection faults	Supported **NOTE** • Supports fiber cut identification. • Supports the identification of a PC connector disconnected from a drop fiber. • Does not support the identification of an APC connector disconnected from a drop fiber.
Demarcation (without a reflector)	Supported **NOTE** • Supports the demarcation for a PC connector connected to a drop fiber. • Does not support the demarcation for an APC connector connected to a drop fiber.
Demarcation (with a reflector)	Supported **NOTE** Does not support E2E attenuation measurement.

2.13 GPON Configuration Guide

GPON configurations include the configurations on GPON profiles, ONTs, and ports. The following section describes configuration methods.

Context

The xPON mode includes two types: distributed (discrete) mode and profile mode. The differences between the two modes are as follows:

- In the distributing mode, ONTs cannot be added in batches. Instead, ONTs need to be configured one by one.
- In profile mode, you can pre-configure ONT line profiles and ONT service profiles and bind ONTs of the same configurations to the same profile to add them in batches, which significantly improves service provisioning efficiency.

The xPON mode is determined during site provisioning and will not be changed.

You can run the **display xpon mode** command to query the xPON mode of the current system.

2.13.1 Configuring a GPON ONT Profile

In distributed mode, GPON ONT profiles include the GPON ONT capability profile and the GPON ONT alarm profile. In profile mode, GPON ONT profiles include DBA profiles, line profiles, service profiles, and alarm profiles. This topic describes how to configure these profiles.

Context

GPON ONT profiles contain the parameters required for configuring the GPON access service, of which,

- DBA profiles specify GPON traffic parameters. The DBA profile bound to an OLT enables the OLT to dynamically allocate bandwidths, improving upstream bandwidth utilization.
- In distributed mode, the GPON ONT capability profile contains the physical port type and quantity of the ONU, mapping mode from service port to GEM port, and traffic control type.
- In profile mode, the line profile is mainly used to configure the information related to DBA, T-CONT, and GEM port. The service profile is used to configure the actual ONT capability and the parameters related to services. The line profile is mandatory and the service profile is optional and dependent of service requirements. Set related attributes in line profile mode and service profile mode, and directly bind the ONT to the line profile and service profile.
- The GPON ONT alarm profile provides a series of alarm threshold parameters that are used for performance measurement and monitoring of activated ONU lines. After a GPON alarm profile is bound to an ONU, the ONU sends alarms to the log host and the NMS if the performance statistics of the line exceed the threshold that is specified in the profile.

NOTE
In this document, ONUs include MDUs and ONTs.

Configuring a DBA Profile

A DBA profile defines the traffic parameters of xPON and can be bound to a T-CONT dynamically allocate the bandwidth and improve the usage of the upstream bandwidth.

Default Configuration

Table 2-12 lists the default settings of the DBA profiles.

Table 2-12 Default settings of the DBA profiles

Parameter	Default Setting	Remarks
Default DBA profile ID in the system	0-9	You can run the **display dba-profile all** command to query the parameter values of each default DBA profile.

Procedure

Add a DBA profile.

Run the **dba-profile add** command to add a DBA profile.

NOTE

- By default, T-CONT is not bound to any DBA profile. Hence, you need to bind a DBA to a T-CONT.
- When you add a DBA profile, the bandwidth value must be a multiple of 64. If you enter a bandwidth value not of a multiple of 64, the system adopts the closest multiple of 64 that is smaller than the value you enter.

Step 1 Query a DBA profile.

Run the **display dba-profile** command to query a DBA profile.

----End

Example

Assume that the name and type of a DBA profile are "DBA_100M" and "type3" respectively, and that the bandwidth required by a user is 100 Mbit/s. To add such a DBA profile, do as follows:

```
huawei(config)#dba-profile add profile-name DBA_100M type3 assure 102400 max 102400
huawei(config)#display dba-profile profile-name DBA_100M
```

Configuring a GPON ONT Capacity Profile (Distributed Mode)

A GPON ONT capability profile identifies the actual capability of a GPON ONU. After an ONT is added and bound to a GPON ONT capability profile, the ONU carries the corresponding services according to parameters configured in the capability profile.

Context

- All GPON ONUs must be bound to the GPON ONT capability profile. Specify the ONT capability profile when running the **ont add** command to add an ONU offline or running the **ont confirm** command to confirm an automatically discovered ONU.
- Currently, the system provides seven default ONT capability profiles that are solidified in the system. The default profiles cannot be modified. The default profile IDs range from 1-7. The reserved ONT capability profile IDs are 8-16.

- The contents of the capability profile restrict the port number that is used in commands for GEM port mapping, T-CONT/PQ mapping, and the ONT VLAN management.
- The ONT capability profile must be configured according to the actual capability of the ONU. Different the capability profile parameters vary according to different ONUs.

Procedure

Run the **ont-profile add** command to configure an ONT capability profile.

- When you add an ONT capability profile, if the profile ID is not specified, the system automatically allocates the least idle profile ID; if the profile name is not specified, the system adopts the default name **ont-profile_x**, where, x is the corresponding ONT capability profile ID.
- The system supports up to 128 ONT capability profiles.
- The system default profiles include the MDU profile and several common ONT (such as OT925, HG850, and HG810) profiles, which can be directly used. It is recommended to manually configure an ONT capability profile only when the default ONT capability profile fails to meet actual requirements.
- When you add an MDU profile manually, the number of the ports must be set to zero.

Step 1 Run the **display ont-profile** command to query the ONT capability profile.

----End

Example

Assume the following parameters: profile ID 30, two POTS ports, four Ethernet ports, mapping mode VLAN ID, and flow control type PQ. To configure such an ONT capability profile for the ONT HG850a and query the capability profile after the configuration is completed, do as follows:

```
huawei(config)#ont-profile add profile-id 30
{ <cr>|profile-name<K> }:

  Command:
         ont-profile add profile-id 30
  Press 'Q' or 'q' to quit input
>  Are you sure you want to set the number of POTS ports to auto-adaptive? (y/n)
 [n]:
>  Number of POTS ports<0-8> [0]:2
>  Are you sure you want to set the number of ETH ports to auto-adaptive? (y/n)
[n]:
>  Number of ETH ports<0-8>  [0]:4
>  Are you sure you want to set the number of VDSL ports to auto-adaptive? (y/n)
[n]: y
>  TDM port type<1-E1,2-T1> [1]:
>  TDM service type<1-TDMoGEM> [1]:
>  Number of TDM ports<0-8> [0]:
>  Number of MOCA ports<0-8> [0]:
>  Are you sure you want to set the number of CATV UNI ports to auto-adaptive? (
y/n) [n]:
>  Number of CATV UNI ports<0-8> [0]:
>  Mapping mode<1-VLANID, 2-802_1pPRI, 3-VLANID_802_1pPRI, 4-PORTID,
   5-PORTID_VLANID, 6-PORTID_802_1pPRI, 7-PORTID_VLANID_802_1pPRI,
```

```
  9-IPTOS, 10-VLANID_IPTOS> [1]:
>  The type of flow control<1-PQ, 2-GEMPORT-CAR, 3-FLOW-CAR> [1]:
  Adding an ONT profile succeeded
  Profile ID  : 30
  Profile name: ont-profile_30

huawei(config)#display ont-profile profile-id 30
  ------------------------------------------------------------------------------
  Profile ID  : 30
  Profile name: ont-profile_30
  ------------------------------------------------------------------------------
  Number of POTS ports:            2
  Number of ETH ports:             4
  Number of VDSL ports:            0
  TDM port type:                  E1
  TDM service type:                TDMoGem
  Number of TDM ports:             0
  Number of MOCA ports:            0
  Number of CATV UNI ports:        0
  Mapping mode:                   VLAN ID
  The type of flow control:        PQ
  ------------------------------------------------------------------------------
  Binding times:                  0
  ------------------------------------------------------------------------------
```

Configuring a GPON ONT Line Profile (Profile Mode)

This topic describes how to configure a GPON ONT line profile and use it when adding an ONT. When an ONT is managed by OMCI or SNMP, the ONT must be bound to a GPON ONT line profile .

Default Configuration

Table 2-13 lists the default settings of a GPON ONT line profile.

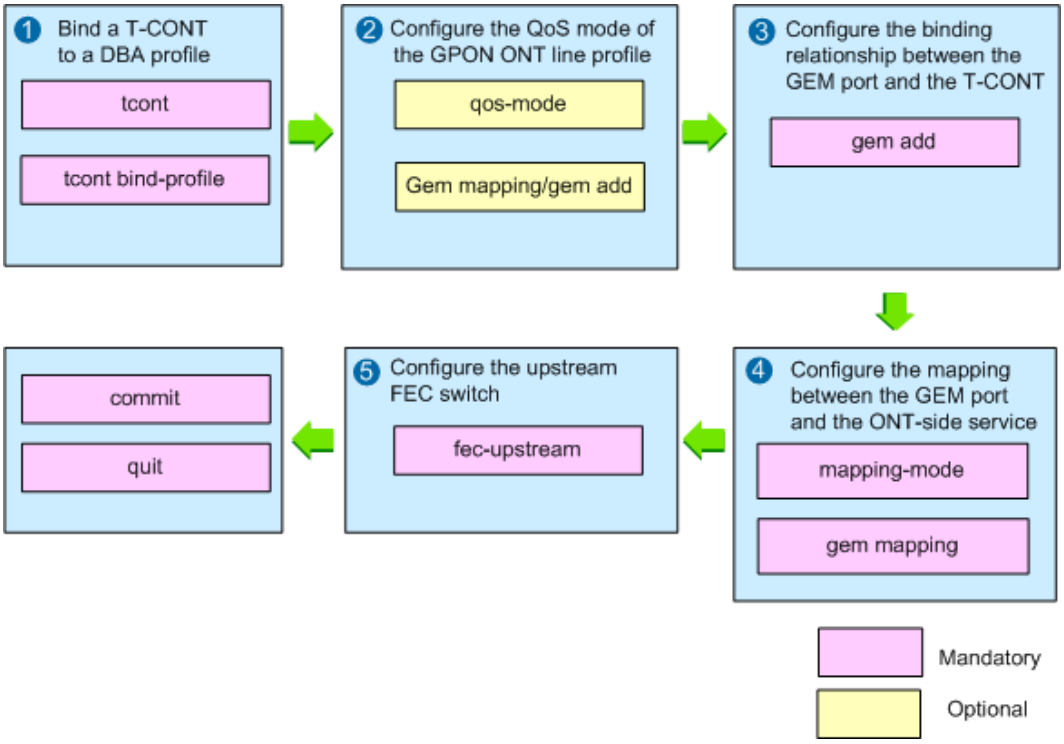
Table 2-13 Default settings of a GPON ONT line profile

Parameter	Default Setting
QoS mode	Priority-queue (PQ) scheduling mode
Mapping mode supported by the ONT	VLAN mapping mode
Upstream FEC switch	Disabled

Configuration Process

Figure 2-54 shows the process of configuring a GPON ONT line profile.

Figure 2-54 Process of configuring a GPON ONT line profile



Procedure

Run the **ont-lineprofile gpon** command to add a GPON ONT line profile and enter the GPON ONT line profile mode.

Regardless of whether the ONT is in the OMCI or SNMP management mode, the line profile must be configured for the ONT. After adding a GPON ONT line profile, directly enter the GPON ONT line profile mode to configure the related attributes of the ONT line.

Step 1 Bind a T-CONT to a DBA profile.

Use the following two methods to bind a DBA profile. Select either method as required. Both methods can coexist in the system.

- In line profile mode:

  This method is applicable to the scenario where the DBA profile is stable and the terminals are of a single type.

  Run the **tcont** command to bind the T-CONT to a DBA profile. Ensure that Configuring a DBA Profile is completed before the configuration.

- In GPON mode:

  This method is applicable to the scenario where the DBA profile changes frequently and the terminals are of different types.

  a. Run the **tcont** command to create a T-CONT, which is not bound to the DBA.

  b. After the configuration of a GPON ONT line profile is complete, enter the GPON mode. Run the **tcont bind-profile** command to bind the T-CONT to a DBA profile. Ensure that Configuring a DBA Profile is completed before the configuration.

By default, T-CONT 0 of an ONT is used by OMCI and is bound to DBA profile 1. The configuration suggestions for the OMCI T-CONT are as follows:

- Do not modify the DBA profile bound to the T-CONT. If you need to modify the profile, ensure that the fixed bandwidth of the modified profile is not lower than 5 Mbit/s.
- Do not bind a GEM port to the T-CONT. That is, ensure that the T-CONT does not carry any service.
- If the sum of the fixed bandwidth and assured bandwidth of the bound DBA profile is larger than the remaining bandwidth of the GPON port, the binding fails and the system displays a message "Failure: The bandwidth is not enough". In this case, you can run the **display port info** command to query the remaining bandwidth (Left guaranteed bandwidth (kbit/s)) of the GPON port, and then decrease the fixed bandwidth and assured bandwidth of the bound DBA profile accordingly.

Step 2 (Optional) Configure the QoS mode of the GPON ONT line profile.

Run the **qos-mode** command to configure the QoS mode of the GPON ONT line profile to be the same as the QoS mode of the GEM port. By default, the QoS mode of the ONT line profile is the PQ scheduling mode. The three QoS modes are as follows:

- flow-car: When this mode is selected, **flow-car** should be selected in the **gem mapping** command, and the maximum traffic depends on the traffic profile bound to the service port. Run the **traffic table ip** command to create a required traffic profile before the configuration.
- gem-car: When this mode is selected, **gem-car** should be selected in the **gem add** command, and the maximum traffic depends on the traffic profile bound to the GEM port.
- priority-queue: When this mode is selected, **priority-queue** should be selected in the **gem add** command. The system has eight default queues (0-7). Queue 7 has the highest priority and the traffic of this queue must be ensured first. The maximum traffic depends on the DBA profile bound to the corresponding T-CONT.

Step 3 Configure the binding relationship between the GEM port and the T-CONT.

Run the **gem add** command to configure the binding relation between the GEM index and the T-CONT in the GPON ONT line profile.

The ONT can carry services only after the mapping between the GEM port and the T-CONT, and the mapping between the GEM port and the service port are configured for the ONT. A correct attribute should be selected for **service-type** based on the service type. Select **eth** when the Ethernet service is carried. Select **tdm** when the TDM service is carried.

Step 4 Configure the mapping between the GEM port and the ONT-side service.

Run the **gem mapping** command to set up the mapping between the GEM port and the ONT-side service.

Before the configuration, run the **mapping-mode** command to configure the mapping mode supported by the ONT to be the same as the configured mapping mode between the GEM port and the ONT-side service. By default, the ONT supports the VLAN mapping mode.

- The mapping modes of the ETH port and the MOCA port are as follows:
  - If the port is specified and then the VLAN is further specified, the mapping mode should be configured to **port-vlan** in the **mapping-mode** command. That is, the port+VLAN mapping mode is used.

  - If the port is specified and then the priority is further specified, the mapping mode should be configured to **port-priority** in the **mapping-mode** command. That is, the port+priority mapping mode is used.
  - If the port and the VLAN are specified and then the priority is further specified, the mapping mode should be configured to **port-vlan-priority** in the **mapping-mode** command. That is, the port+VLAN+priority mapping mode is used.
- As a special port, the IPHOST or E1 port is not restricted by the ONT mapping mode.

When the mapping mode is **vlan-priority** or **port-vlan-priority**,

- If a GEM port is mapped to multiple VLANs, any of these VLANs cannot map to any other GEM port.
- If a VLAN is mapped to multiple GEM ports, any of these GEM ports cannot map to any other VLAN.

Step 5 Configure the upstream FEC switch.

Run the **fec-upstream** command to configure the upstream FEC switch of the GPON ONT line profile. By default, this switch is disabled.

In the FEC check, the system inserts redundancy data into normal packets. In this way, the line has certain error tolerant function, but certain bandwidth resources are wasted. Enabling the FEC function enhances the error tolerant capability of the line but occupies certain bandwidth. Therefore, determine whether to enable the FEC function based on the actual line planning.

Step 6 Run the **commit** command to make the parameters of the profile take effect. The configuration of a line profile takes effect only after you perform this operation.

NOTE
If this profile is not bound, all the parameters that are configured take effect when the profile is bound. If this profile is already bound, the configuration takes effect on all ONTs bound to this profile immediately.

Step 7 Run the **quit** command to return to the global configuration mode.

----End

Example

Assume that the GEM index is 1, the GEM port is bound to T-CONT 1 and mapped to ETH 1 of the ONT. To add GPON ONT line profile 5, create a channel for carrying the Ethernet service, with T-CONT 1 and bound to DBA profile 12, use the QoS policy of controlling the traffic based on GEM ports, and bind the GEM port to default traffic profile 6, do as follows:

```
huawei(config)#ont-lineprofile gpon profile-id 5
huawei(config-gpon-lineprofile-5)#tcont 1 dba-profile-id 12
huawei(config-gpon-lineprofile-5)#qos-mode gem-car
huawei(config-gpon-lineprofile-5)#gem add 1 eth tcont 1 gem-car 6
huawei(config-gpon-lineprofile-5)#mapping-mode port
huawei(config-gpon-lineprofile-5)#gem mapping 1 0 eth 1
huawei(config-gpon-lineprofile-5)#commit
huawei(config-gpon-lineprofile-5)#quit
```

To modify GPON ONT line profile 5, and change the DBA profile bound to T-CONT 1 from DBA profile 12 to DBA profile 10, do as follows:

```
huawei(config)#ont-lineprofile gpon profile-id 5
huawei(config-gpon-lineprofile-5)#tcont 1 dba-profile-id 10
huawei(config-gpon-lineprofile-5)#commit
huawei(config-gpon-lineprofile-5)#quit
```

To modify GPON ONT line profile 5, bind GEM index 1 to T-CONT 2, and map GEM index 1 to ONT ETH port 2, do as follows:

NOTE

If a GEM index is used by a traffic stream, delete this traffic stream first and then the GEM index.

```
huawei(config)#ont-lineprofile gpon profile-id 5
huawei(config-gpon-lineprofile-5)#undo gem mapping 1 0
huawei(config-gpon-lineprofile-5)#gem delete 1
huawei(config-gpon-lineprofile-5)#gem add 1 eth tcont 2
huawei(config-gpon-lineprofile-5)#gem mapping 1 0 eth 2
huawei(config-gpon-lineprofile-5)#commit
huawei(config-gpon-lineprofile-5)#quit
```

Configuring a GPON ONT Service Profile

The GPON ONT service profile provides a channel for configuring the service of the ONT managed in the OMCI mode. The ONT (such as the MDU) managed in the SNMP mode does not suppport the configuration of the GPON ONT service profile. To configure the service of the ONT (such as the MDU) managed in the SNMP mode, you need to log in to the ONT.

Default Configuration

Table 2-14 lists the default settings of the GPON ONT service profile.

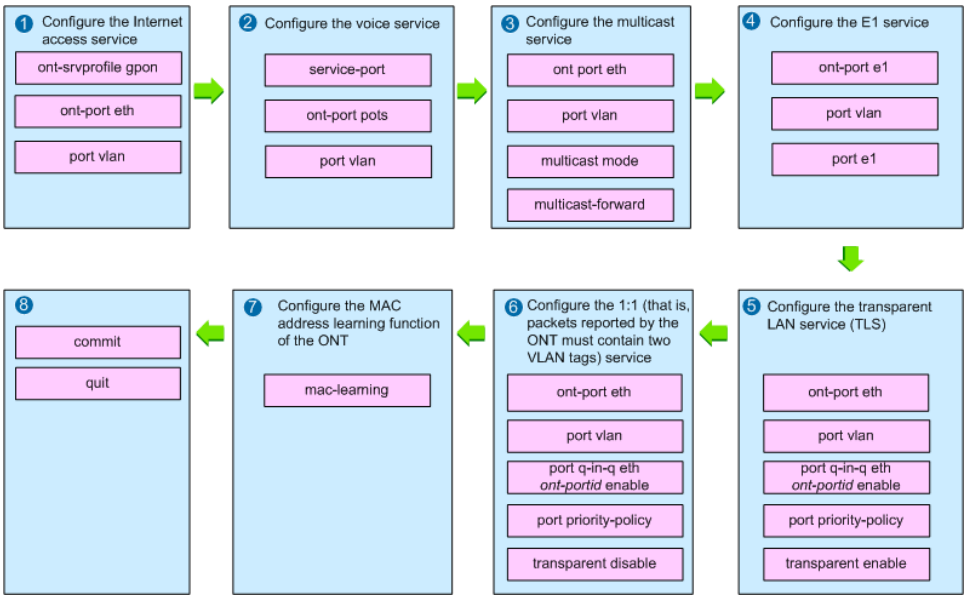
Table 2-14 Default settings of the GPON ONT service profile

Parameter	Default Setting
Multicast mode of the ONT	Unconcern (the OLT does not perform any processing)
Mode for the ONT to process the VLAN tag of the multicast data packets	Unconcern
Coding mode for the E1 port of the ONT	HDB3
Source of the priority copied for the upstream packets on the ONT port	Unconcern
QinQ attribute for the Ethernet port of the ONT	Unconcern
Transparent transmission function of the ONT	Disabled
MAC address learning function of the ONT	Enabled

Configuration Process

Figure 2-55 shows the process of configuring a GPON ONT service profile.

Figure 2-55 Process of configuring a GPON ONT service profile



Procedure

Run the **ont-srvprofile gpon** command to add a GPON ONT service profile, and then enter the GPON ONT service profile mode.

If the ONT management mode is the SNMP mode, you do not need to configure the service profile. After adding a GPON ONT service profile, directly enter the GPON ONT service profile mode to configure the related items. Select the configuration items according to the service requirements.

Step 1 Configure the Internet access service.

1. Run the **ont-port eth** command to configure the port capability set of the ONT. The capability set plans various types of ports supported by the ONT. The port capability set in the ONT service profile must be the same as the actual ONT capability set.

   If the port capability set of an ONT is set to **adaptive**, the OLT automatically adapts to the online ONT according to the OLT's actual capability. By default, eight ETH ports and one IPHOST are displayed.

2. Run the **port vlan** command to configure the port VLAN of the ONT.

Step 2 Configure the voice service.

NOTE

The voice service of the ONT is configured by issuing an XML file to the NMS and the OLT performs only transparent transmission. You only need to run the **service-port** command to create a service port carrying the voice service.

1. Run the **ont-port pots** command to configure the port capability set of the ONT. The port capability set in the ONT service profile must be the same as the actual ONT capability set.

2. Run the **port vlan** command to configure the port VLAN of the ONT.

Step 3 Configure the multicast service.

1. Run the **ont-port eth** command to configure the port capability set of the ONT. The port capability set in the ONT service profile must be the same as the actual ONT capability set.

   If the port capability set of an ONT is set to **adaptive**, the OLT automatically adapts to the online ONT according to the OLT's actual capability. By default, eight ETH ports and one IPHOST are displayed.
2. Run the **port vlan** command to configure the port VLAN of the ONT.
3. Run the **multicast mode** command to configure the multicast mode of the ONT. By default, the multicast mode of the ONT is **unconcern**.
   - Unconcern: indicates the unconcern mode. After this mode is selected, the OLT does not limit the multicast mode, and the multicast mode on the OLT automatically matches the multicast mode on the ONT.
   - Igmp-snooping: IGMP snooping obtains the related information and maintains the multicast forwarding entries by listening to the IGMP packets in the communication between the user and the multicast router.
   - Olt-control: indicates the dynamic controllable multicast mode. A multicast forwarding entry can be created for the multicast join packet of the user only after the packet passes the authentication. This mode is supported by the MDU, but is not supported by the ONT.
4. Run the **multicast-forward** command to configure the processing mode on the VLAN tag of the multicast data packets for the ONT. By default, the multicast forwarding mode of the ONT is **unconcern**.
   - Unconcern: indicates the unconcern forwarding mode. After this mode is selected, the OLT does not process the VLAN tag of the multicast data packets.
   - Tag: Set the multicast forwarding mode to contain the VLAN tag. To transparently transmit the VLAN tag of the multicast packets, select **transparent**. To switch the VLAN tag of the multicast packets, select **translation**, and then configure the VLAN ID that is switched to.
   - Untag: Set the multicast forwarding mode not to contain the VLAN tag.

Step 4 Configure the E1 service.

1. Run the **ont-port e1** command to configure the port capability set of the ONT. The port capability set in the ONT service profile must be the same as the actual ONT capability set.
2. Run the **port vlan** command to configure the port VLAN of the ONT.
3. Run the **port e1** command to configure the coding mode supported by the E1 port of the ONT. By default, the E1 port supports the HDB3 coding mode. The coding mode must be the same as that on the interconnected device.

Step 5 Configure the transparent LAN service (TLS).

1. Run the **ont-port eth** command to configure the port capability set of the ONT. The port capability set in the ONT service profile must be the same as the actual ONT capability set.

   If the port capability set of an ONT is set to **adaptive**, the OLT automatically adapts to the online ONT according to the OLT's actual capability. By default, eight ETH ports and one IPHOST are displayed.
2. Run the **port vlan** command to configure the port VLAN of the ONT.

3. Run the **port q-in-q eth *ont-portid* enable** command to enable the QinQ function of the Ethernet port on the ONT. By default, the QinQ function of the Ethernet port on the ONT is unconcerned.
4. Run the **port priority-policy** command to configure the source of the priority copied for the upstream packets on the ONT port. By default, the source of the priority copied for the upstream packets on the ONT Ethernet port is unconcerned.
   - Unconcern: The source of the priority copied for the upstream packets on the Ethernet port of the ONT is not concerned.
   - assigned: Specifies the priority. Run the **ont port native-vlan** command to specify the **priority** of the port.
   - Copy-cos: Copy the priority. Copy the priority from C-TAG.
5. Run the **transparent enable** command to enable the transparent transmission function of the ONT. By default, the transparent transmission function of the ONT is disabled. After the transparent transmission function of the ONT is enabled, all packets (including service packets and protocol packets) are transparently transmitted by the ONT.

NOTE
The service port for the TLS service must also be of the TLS type. Run the **service-port** command to create a service port of the TLS type. Select **other-all** for the multi-service type.

Step 6 Configure the 1:1 (that is, packets reported by the ONT must contain two VLAN tags) service.

1. Run the **ont-port eth** command to configure the port capability set of the ONT. The port capability set in the ONT service profile must be the same as the actual ONT capability set.

   If the port capability set of an ONT is set to **adaptive**, the OLT automatically adapts to the online ONT according to the OLT's actual capability. By default, eight ETH ports and one IPHOST are displayed.
2. Run the **port vlan** command to configure the port VLAN of the ONT.
3. Run the **port q-in-q eth *ont-portid* enable** command to enable the QinQ function of the Ethernet port on the ONT. By default, the QinQ function of the Ethernet port on the ONT is unconcerned.
4. Run the **port priority-policy** command to configure the source of the priority copied for the upstream packets on the ONT port. By default, the source of the priority copied for the upstream packets on the ONT Ethernet port is unconcerned.
   - Unconcern: The source of the priority copied for the upstream packets on the Ethernet port of the ONT is not concerned.
   - assigned: Specifies the priority. Run the **ont port native-vlan** command to specify the **priority** of the port.
   - Copy-cos: Copy the priority. Copy the priority from C-TAG.
5. Run the **transparent disable** command to disable the transparent transmission function of the ONT.

Step 7 Run the **mac-learning** command to configure the MAC address learning function of the ONT. This function is enabled by default.

Step 8 Run the **commit** command to make the parameters of the profile take effect. The configuration of the service profile takes effect only after you perform this operation.

NOTE
If this profile is not bound, all the parameters that are configured take effect when the profile is bound. If this profile is already bound, the configuration takes effect on all ONTs bound to this profile immediately.

Step 9 Run the **quit** command to return to the global config mode.

----End

Example

Assume that the profile is used for the Internet access service, the ONT supports four ETH ports, and the VLAN ID of the ETH ports is 10. To add GPON ONT service profile 5, do as follows:

```
huawei(config)#ont-srvprofile gpon profile-id 5
huawei(config-gpon-srvprofile-5)#ont-port eth adaptive
huawei(config-gpon-srvprofile-5)#port vlan eth 1-4 10
huawei(config-gpon-srvprofile-5)#commit
huawei(config-gpon-srvprofile-5)#quit
```

Assume that the profile is used for the multicast service, the ONT supports four ETH ports, the VLAN ID of the ETH ports is 100, and the multicast mode of the ONT is the controllable multicast mode (you need to switch the multicast VLAN tag to 841 because the STB only supports carrying the VLAN tag of 841). To add GPON ONT service profile 6, do as follows:

```
huawei(config)#ont-srvprofile gpon profile-id 6
huawei(config-gpon-srvprofile-6)#ont-port eth adaptive
huawei(config-gpon-srvprofile-6)#port vlan eth 1-4 100
huawei(config-gpon-srvprofile-6)#multicast mode olt-control
huawei(config-gpon-srvprofile-6)#multicast-forward tag translation 841
huawei(config-gpon-srvprofile-6)#commit
huawei(config-gpon-srvprofile-6)#quit
```

Configuring a GPON ONT Alarm Profile

This topic describes how to add an alarm profile, and configure most of the performance parameters for various ONT lines as a profile. After the alarm profile is configured and bound successfully, the ONT can directly use the profile when it is activated.

Context

An ONT alarm profile defines a series of alarm thresholds that are used to monitor the performance of an activated ONT line. When the statistics result of a parameter reaches the alarm threshold, the NE is notified and an alarm is sent to the log server and the NMS.

- The MA5600T/MA5603T/MA5608T supports up to 50 alarm profiles.
- The system contains a default alarm profile with the ID 1. This profile cannot be deleted but can be modified.

Procedure

Run the **gpon alarm-profile add** command to add a GPON ONT alarm profile.

All parameters in the default profile are set to 0, which indicates that no alarm is reported. When an alarm profile is created, the default values of all alarm thresholds are 0, which indicates that no alarm is reported.

Step 1 Run the **display gpon alarm-profile** command to query the alarm profile.

----End

Example

To add GPON ONT alarm profile 5, set the alarm threshold for the packet loss of the GEM port to 10, set the alarm threshold for the number of mis-transmitted packets to 30, and use the default value 0 for all other thresholds, do as follows:

```
huawei(config)#gpon alarm-profile add profile-id 5
{ <cr>|profile-name<K> }:

  Command:
         gpon alarm-profile add profile-id 5
  Press 'Q' or 'q' to quit input
>  GEM port loss of packets threshold (0~100)[0]:50
>  GEM port misinserted packets threshold (0~100)[0]:50
>  GEM port impaired blocks threshold (0~100)0[0]:50
>  Ethernet FCS errors threshold (0~100)[0]:50
>  Ethernet excessive collision count threshold (0~100)[0]:50
>  Ethernet late collision count threshold (0~100)[0]:50
>  Too long Ethernet frames threshold (0~100)[0]:50
>  Ethernet buffer (Rx) overflows threshold (0~100)[0]:50
>  Ethernet buffer (Tx) overflows threshold (0~100)[0]:50
>  Ethernet single collision frame count threshold (0~100)[0]:50
>  Ethernet multiple collisions frame count threshold (0~100)[0]:50
>  Ethernet SQE count threshold (0~100)[0]:50
>  Ethernet deferred transmission count threshold (0~100)[0]:50
>  Ethernet internal MAC Tx errors threshold (0~100)[0]:50
>  Ethernet carrier sense errors threshold (0~100)[0]:50
>  Ethernet alignment errors threshold (0~100)[0]:50
>  Ethernet internal MAC Rx errors threshold (0~100)[0]:50
>  PPPOE filtered frames threshold (0~100)[0]:50
>  MAC bridge port discarded frames due to delay threshold (0~100)[0]:50
>  MAC bridge port MTU exceeded discard frames threshold (0~100)[0]:50
>  MAC bridge port received incorrect frames threshold (0~100)[0]:50
>  CES general error time threshold(0~100)[0]:
>  CES severely time threshold(0~100)[0]:
>  CES bursty time threshold(0~100)[0]:
>  CES controlled slip threshold(0~100)[0]:
>  CES unavailable time threshold(0~100)[0]:
>  Drop events threshold(0~100)[0]:
>  Undersize packets threshold(0~100)[0]:
>  Fragments threshold(0~100)[0]:
>  Jabbers threshold(0~100)[0]:
>  Failed signal of ONT threshold(Format:1e-x, x: 3~8)[3]:5
>  Degraded signal of ONT threshold(Format:1e-x, x: 4~9)[4]:6
>  FEC uncorrectable code words threshold(0~1101600000)[0]:6
>  FEC correctable code words threshold(0~1101600000)[0]:6
>  Upstream PQ discarded byte alarm threshold(0~65535)[0]:6
>  Downstream PQ discarded byte alarm threshold(0~65535)[0]:6
>  Encryption key errors threshold(0~100)[0]:6
>  XGEM key errors threshold(0~100)[0]:6
>  XGEM HEC error count threshold(0~100)[0]:6

  Adding an alarm profile succeeded
```

```
  Profile ID  : 5
  Profile name: alarm-profile_5

huawei(config)#display gpon alarm-profile profile-id 5
  ------------------------------------------------------------
  Profile ID  : 5
  Profile name: alarm-profile_5
  ------------------------------------------------------------
  GEM port loss of packets threshold:                    50
  GEM port misinserted packets threshold:                 50
  GEM port impaired blocks threshold:                    50
  Ethernet FCS errors threshold:                         50
  Ethernet excessive collision count threshold:           50
  Ethernet late collision count threshold:                50
  Too long Ethernet frames threshold:                    50
  Ethernet buffer (Rx) overflows threshold:               50
  Ethernet buffer (Tx) overflows threshold:               50
  Ethernet single collision frame count threshold:         50
  Ethernet multiple collisions frame count threshold:      50
  Ethernet SQE count threshold:                          50
  Ethernet deferred transmission count threshold:          50
  Ethernet internal MAC Tx errors threshold:              50
  Ethernet carrier sense errors threshold:                50
  Ethernet alignment errors threshold:                   50
  Ethernet internal MAC Rx errors threshold:              50
  PPPOE filtered frames threshold:                       50
  MAC bridge port discarded frames due to delay threshold:  50
  MAC bridge port MTU exceeded discard frames threshold:    50
  MAC bridge port received incorrect frames threshold:     50
  CES general error time threshold:                       0
  CES severely time threshold:                           0
  CES bursty time threshold:                             0
  CES controlled slip time threshold:                     0
  CES unavailable time threshold:                        0
  Drop events threshold:                                 0
  Undersize packets threshold:                           0
  Fragments threshold:                                   0
  Jabbers threshold:                                     0
  Failed signal of ONU threshold (Format:1e-x):           5
  Degraded signal of ONU threshold (Format:1e-x):          6
  FEC uncorrectable code words threshold:                 6
  FEC correctable code words threshold:                   6
  Upstream PQ discarded byte alarm threshold:             6
  Downstream PQ discarded byte alarm threshold:           6
  Encryption key errors threshold:                       6
  XGEM key errors threshold:                             6
  XGEM HEC error count threshold:                         6
  ------------------------------------------------------------
  Binding Times:                                        0
  ------------------------------------------------------------
```

2.13.2 Configuring a GPON ONT (Distributed Mode)

The MA5600T/MA5603T/MA5608T provides end users with services through the ONT. The MA5600T/MA5603T/MA5608T can manage the ONT and the ONT can work in the normal

state only after the channel between the MA5600T/MA5603T/MA5608T and the ONT is available.

Prerequisites

The GPON ONT profile is already created. Configuring a GPON ONT Capacity Profile (Distributed Mode) and Configuring a GPON ONT Alarm Profile are already completed.

Context

The MA5600T/MA5603T/MA5608T uses the ONT Management and Control Interface (OMCI) protocol to manage and configure the GPON ONT, and supports the offline configuration of the ONT. The ONT does not need to save the configuration information locally. This helps to provision services.

Table 2-15 lists the default settings of the GPON ONT.

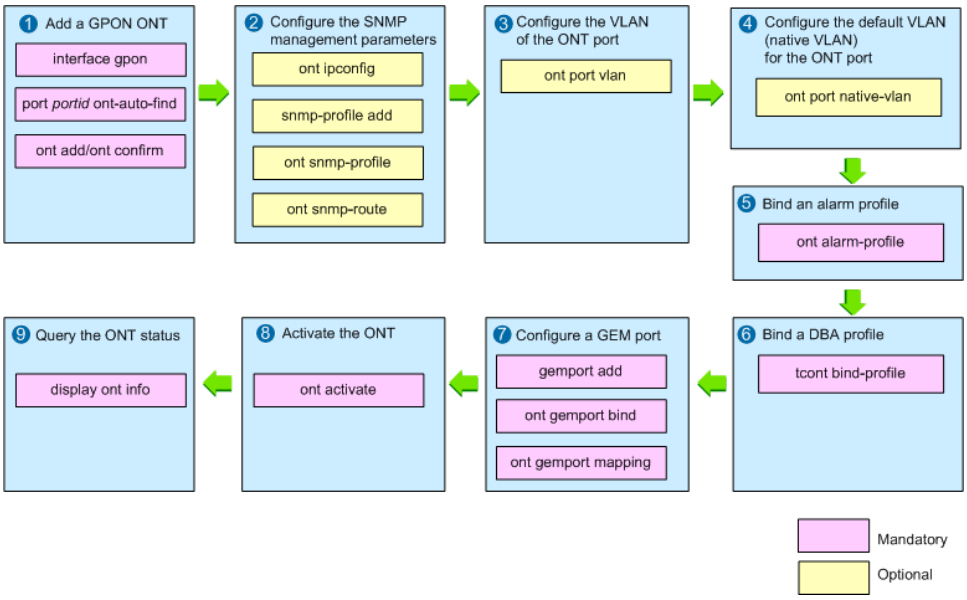
Table 2-15 Default settings of the GPON ONT

Parameter	Default Setting
ONT auto-find function of a GPON port	Disabled
ONT status after an ONT is added	Activated
Default VLAN of the ONT port	1

Configuration Process

Figure 2-56 shows the process of configuring a GPON ONT.

Figure 2-56 Process of configuring a GPON ONT



Procedure

Add a GPON ONT.

1. Run the **interface gpon** command to enter the GPON mode.
2. Run the **port *portid* ont-auto-find** command to enable the auto-find function of the ONT. After the function is enabled, you can add an ONT according to the information reported by the system. By default, the ONT auto-find function of a GPON port is disabled.

   NOTE

   An auto-find ONT is in the auto-find state. The auto-find ONT can work in the normal state only after it is confirmed or added.
3. Run the **ont add** command to add an ONT offline, or run the **ont confirm** command to confirm the auto-find ONT.

   When ONTs are added or confirmed, the system provides four authentication modes: SN, password, SN+password, LOID+CHECKCODE.

   - SN authentication: The OLT detects the serial number (SN) reported by an ONT. If the SN is consistent with the OLT configuration, authentication is passed and the ONT goes online. This mode requires recording all ONT SNs. Hence, it is used to confirm auto discovery ONTs and is not applicable to adding ONTs in batches.
   - Password authentication: The OLT detects the password reported by an ONT. If the password is consistent with the OLT configuration, the ONT goes online normally. This mode requires planning ONT passwords and does not require manually recording ONT SNs. Hence, it is applicable to adding ONTs in batches. The password authentication provides two discovery modes: always-on and once-on.
     - always-on: After first password authentication is passed, no SN is allocated and password authentication is always used in subsequent authentications. This discovery mode is easy for future maintenance. In the always-on discovery mode, configuration is not required to be modified when an ONT is replaced and only the password is required. The always-on discovery mode has lower security. If other users know the password, the users will illegally have service permissions.
     - Once-on: After first password authentication is passed, an SN is automatically allocated and password+SN authentication is used in subsequent authentications. An ONT can go online only after the correct password and SN are entered. The once-on authentication mode has high security. After an ONT is replaced or the password is mistakenly changed, the ONT needs to be configured again, which requires more maintenance effort.
   - SN+password authentication: The OLT detects the password and SN reported by an ONT. If the password and SN are consistent with the OLT configuration, the ONT goes online normally. This authentication mode has the highest security but it requires manually recording ONT SNs.
   - LOID+CHECKCODE authentication: defined by a telecom operator. In this authentication mode, LOID has 24 bytes, and CHECKCODE has 12 bytes and is optional. Whether 24 bytes or 36 bytes are used for authentication depends on data planning, which is unified over the entire network. The OLT determines whether LOID+CHECKCODE reported by the ONT is the same as the configured one. If they are the same, the ONT authentication is passed. If they are different, the OLT obtains the ONT password and compares it with the last 10 bytes of the LOID. If they are the same, the ONT authentication is also passed. This operation is for compatibility with the ONTs using password authentication.

   Adding ONTs in offline mode is applicable to the batch deployment scenario. All ONTs are added to the OLT to complete service provisioning beforehand. When a use

subscribes to the service, an installation engineer takes an ONT to the user's house and completes configurations. After the ONT goes online and passes authentication (generally the password authentication mode or LOID authentication mode is used), the service is provisioned.

Adding ONTs in auto discovery mode is applicable to the scenario where a small number of ONTs are added. When users subscribe to the service, installation engineers take ONTs to the users' houses. After the ONTs go online, the OLT confirms the ONTs one by one. Generally, the MAC address authentication mode is used to confirm the ONTs.

NOTE

- If the ONU is an independent NE and is directly managed by the NMS through the SNMP management mode, select the SNMP management mode. For this mode, you only need to configure the parameters for the GPON line and the parameters for the management channel on the OLT.
- If the ONU is not an independent NE and all its configuration data is issued by the OLT through OMCI, select the OMCI management mode. For this mode, you need to configure all parameters (including line parameters, UNI port parameters, and service parameters) that are required for the ONU on the OLT.
- Generally, the ONT management mode is set to the OMCI mode.

4. (Optional) When the ONT management mode is the SNMP mode, you need to configure the SNMP management parameters for the ONT. The procedure is as follows:
   a. Run the **ont ipconfig** command to configure the management IP address of the ONT.

      The IP address should not be in the same subnet for the IP address of the VLAN port.
   b. Run the **ont snmp-profile** command to bind the ONT with an SNMP profile.

      Run the **snmp-profile add** command to add an SNMP profile before the configuration.
   c. Run the **ont snmp-route** command to configure a static route for the NMS server, that is, configure the IP address of the next hop.

Step 1 (Optional) Configure the VLAN of the ONT port.

Run the **ont port vlan** command to configure the VLAN of the ONT port. By default, all the ports on the ONT belong to VLAN 1.

Step 2 (Optional) Configure the default VLAN (native VLAN) for the ONT port.

Run the **ont port native-vlan** command to configure the default VLAN for the ONT port. By default, the default VLAN ID of the ONT port is 1.

- If the packets reported from a user (such a PC) to the ONT are untagged, the packets are tagged with the default VLAN of the port on the ONT and then reported to the OLT.
- If the packets reported from a user to the ONT are tagged, you need to configure the port VLAN of the ONT to be the same as the VLAN in the user tag. The packets are not tagged with the default VLAN of the port on the ONT but are reported to the OLT with the user tag.

Step 3 Bind an alarm profile.

Run the **ont alarm-profile** command to bind an alarm profile. Ensure that Configuring a GPON ONT Alarm Profile is completed before the configuration.

Step 4 Bind a DBA profile.

Run the **tcont bind-profile** command to bind a DBA profile to a T-CONT.

A DBA profile can be bound to a T-CONT after an ONT is added.

Step 5 Configure a GEM port.

1. Run the **gemport add** command to add a GEM port. When adding a GEM port, select the correct attribute according to the service type.
2. Run the **ont gemport bind** command to bind the GEM port to an ONT T-CONT, that is, allocating the T-CONT resources to the GEM port.

NOTE
If traffic streams are configured on a GEM port and an ONT is the working ONT in a single-homing protection group, the GEM port cannot be bound to or unbound from the ONT.

3. Run the **ont gemport mapping** command to create the mapping between the GEM port and the ONT-side service.

Step 6 Activate the ONT.

Run the **ont activate** command to activate the ONT. The ONT can transmit services only when it is in the activated state.

After being added, the ONT is in the activated state by default. The step is required only when the ONT is in the deactivated state.

Step 7 Query the ONT status.

Run the **display ont info** command to query the ONT running status, configuration status, and matching status.

----End

Example

To add five ONTs in offline mode with password authentication mode (ONT passwords are 0100000001-0100000005), set the discovery mode of password authentication to always-on, and bind ONT capability profile 10, do as follows:

```
huawei(config)#interface gpon 0/2
huawei(config-if-gpon-0/2)#ont add 0 password-auth 0100000001 always-on profile-id 10
manage-mode omci
huawei(config-if-gpon-0/2)#ont add 0 password-auth 0100000002 always-on profile-id 10
manage-mode omci
huawei(config-if-gpon-0/2)#ont add 0 password-auth 0100000003 always-on profile-id 10
manage-mode omci
huawei(config-if-gpon-0/2)#ont add 0 password-auth 0100000004 always-on profile-id 10
manage-mode omci
huawei(config-if-gpon-0/2)#ont add 0 password-auth 0100000005 always-on profile-id 10
manage-mode omci
```

To add an ONT that is managed by the OLT through the OMCI protocol, confirm this ONT according to the SN 3230313185885B41 automatically reported by the system, and bind the ONT with capability profile 3 that match the ONT, do as follows:

```
huawei(config)#interface gpon 0/2
huawei(config-if-gpon-0/2)#port 0 ont-auto-find enable
huawei(config-if-gpon-0/2)#ont confirm 0 sn-auth 3230313185885B41 profile-id 3
manage-mode omci
```

To add an ONU that is managed as an independent NE and whose SN is known as 3230313185885641, bind the ONU with capability profile 4 that matches the ONU, configure the NMS parameters for the ONU, and set the management VLAN to 100, do as follows:

```
huawei(config)#snmp-profile add profile-id 1 v2c public private 10.10.5.53 161 huawei
huawei(config)#interface gpon 0/2
huawei(config-if-gpon-0/2)#ont add 0 2 sn-auth 3230313185885641 profile-id 4
manage-mode snmp
huawei(config-if-gpon-0/2)#ont ipconfig 0 2 static ip-address 10.20.20.20 mask
255.255.255.0 gateway 10.10.20.1 vlan 100
huawei(config-if-gpon-0/2)#ont snmp-profile 0 2 profile-id 1
huawei(config-if-gpon-0/2)#ont snmp-route 0 2 ip-address 10.10.20.190 mask
255.255.255.0 next-hop 10.10.20.100
```

2.13.3 Configuring a GPON ONT (Profile Mode)

The MA5600T/MA5603T/MA5608T provides end users with services through the ONT. The MA5600T/MA5603T/MA5608T can manage the ONT and the ONT can work in the normal state only after the channel between the MA5600T/MA5603T/MA5608T and the ONT is available.

Prerequisites

The GPON ONT profile is already created.

- For an ONT, Configuring a GPON ONT Line Profile (Profile Mode), Configuring a GPON ONT Service Profile, and Configuring a GPON ONT Alarm Profile are already completed.
- For an MDU or ONU, Configuring a GPON ONT Line Profile (Profile Mode) and Configuring a GPON ONT Alarm Profile are already completed.

Context

The MA5600T/MA5603T/MA5608T uses the ONT Management and Control Interface (OMCI) protocol to manage and configure the GPON ONT, and supports the offline configuration of the ONT.

In the profile mode, the related configuration of the GPON ONT is already integrated in the service profile and the line profile. When adding an ONT, you only need to bind the ONT with the corresponding service profile and line profile.

Table 2-16 lists the default settings of the GPON ONT.

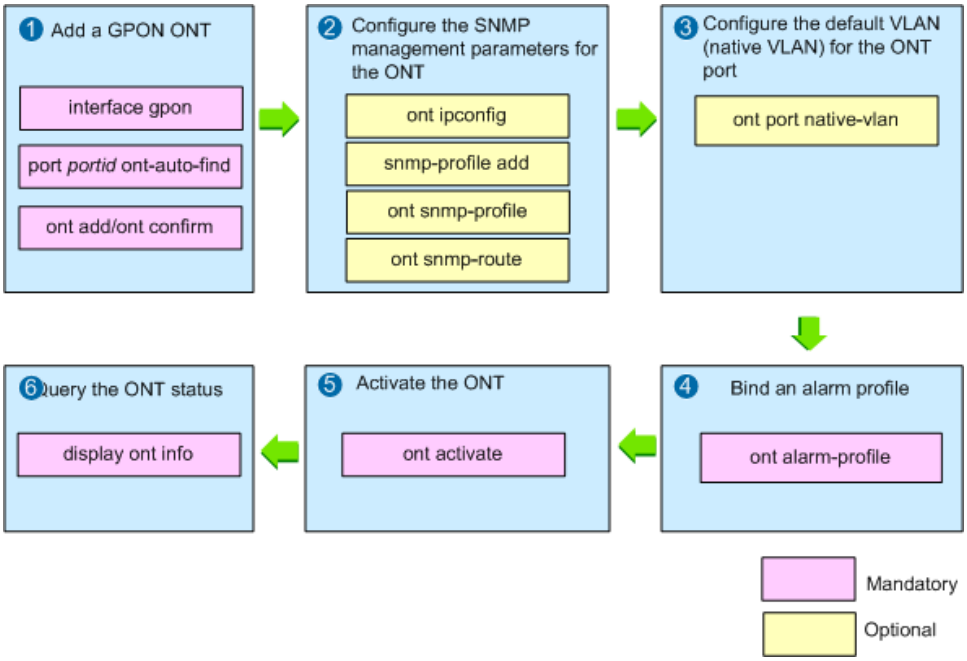
Table 2-16 Default settings of the GPON ONT

Parameter	Default Setting
ONT auto-find function of a GPON port	Disabled
ONT status after an ONT is added	Activated
Default VLAN of the ONT port	1

Configuration Process

Figure 2-57 shows the process of configuring a GPON ONT.

Figure 2-57 Process of configuring a GPON ONT



Procedure

Run the **interface gpon** command to enter the GPON mode.

Step 1 Add a GPON ONT.

1. Run the **port *portid* ont-auto-find** command to enable the auto discovery function of the ONT. After the function is enabled, you can add an ONT according to the information reported by the system. By default, the ONT auto discovery function of a GPON port is disabled.

NOTE

An auto discovery ONT is in the auto discovery state. The auto discovery ONT can work in the normal state only after it is confirmed or added.

2. Run the **ont add** command to add an ONT offline, or run the **ont confirm** command to confirm the auto discovery ONT.

   When ONTs are added or confirmed, the system provides four authentication modes: SN, password, SN+password, LOID+CHECKCODE.

   - SN authentication: The OLT detects the serial number (SN) reported by an ONT. If the SN is consistent with the OLT configuration, authentication is passed and the ONT goes online. This mode requires recording all ONT SNs. Hence, it is used to confirm auto discovery ONTs and is not applicable to adding ONTs in batches.
   - Password authentication: The OLT detects the password reported by an ONT. If the password is consistent with the OLT configuration, the ONT goes online normally. This mode requires planning ONT passwords and does not require manually recording ONT SNs. Hence, it is applicable to adding ONTs in batches. The password authentication provides two discovery modes: always-on and once-on.
     - always-on: After first password authentication is passed, no SN is allocated and password authentication is always used in subsequent authentications. This discovery mode is easy for future maintenance. In the always-on discovery

mode, configuration is not required to be modified when an ONT is replaced and only the password is required. The always-on discovery mode has lower security. If other users know the password, the users will illegally have service permissions.

- Once-on: After first password authentication is passed, an SN is automatically allocated and password+SN authentication is used in subsequent authentications. An ONT can go online only after the correct password and SN are entered. The once-on authentication mode has high security. After an ONT is replaced or the password is mistakenly changed, the ONT needs to be configured again, which requires more maintenance effort.

- SN+password authentication: The OLT detects the password and SN reported by an ONT. If the password and SN are consistent with the OLT configuration, the ONT goes online normally. This authentication mode has the highest security but it requires manually recording ONT SNs.
- LOID+CHECKCODE authentication: defined by a telecom operator. In this authentication mode, LOID has 24 bytes, and CHECKCODE has 12 bytes and is optional. Whether 24 bytes or 36 bytes are used for authentication depends on data planning, which is unified over the entire network. The OLT determines whether LOID+CHECKCODE reported by the ONT is the same as the configured one. If they are the same, the ONT authentication is passed. If they are different, the OLT obtains the ONT password and compares it with the last 10 bytes of the LOID. If they are the same, the ONT authentication is also passed. This operation is for compatibility with the ONTs using password authentication.

Adding ONTs in offline mode is applicable to the batch deployment scenario. All ONTs are added to the OLT to complete service provisioning beforehand. When a use subscribes to the service, an installation engineer takes an ONT to the user's house and completes configurations. After the ONT goes online and passes authentication (generally the password authentication mode or LOID authentication mode is used), the service is provisioned.

Adding ONTs in auto discovery mode is applicable to the scenario where a small number of ONTs are added. When users subscribe to the service, installation engineers take ONTs to the users' houses. After the ONTs go online, the OLT confirms the ONTs one by one. Generally, the MAC address authentication mode is used to confirm the ONTs.

NOTE

- If the ONU is an independent NE and is directly managed by the NMS through the SNMP management mode, select the SNMP management mode. For this mode, you only need to configure the parameters for the GPON line and the parameters for the management channel on the OLT. You only need to bind the ONU with a line profile.
- If the ONU is not an independent NE and all its configuration data is issued by the OLT through OMCI, select the OMCI management mode. For this mode, you need to configure all parameters (including line parameters, UNI port parameters, and service parameters) that are required for the ONU on the OLT. Configuring management channel parameters is not supported. You need to bind the ONT with a line profile and a service profile.
- Generally, the ONT management mode is set to the OMCI mode. You need to bind the ONT with a line profile and a service profile.

3. (Optional) When the ONT management mode is the SNMP mode, you need to configure the SNMP management parameters for the ONT. The procedure is as follows:
   a. Run the **ont ipconfig** command to configure the management IP address of the ONT.

      The IP address should not be in the same subnet for the IP address of the VLAN port.

   b. Run the **ont snmp-profile** command to bind the ONT with an SNMP profile.

Run the **snmp-profile add** command to add an SNMP profile before the configuration.

c. Run the **ont snmp-route** command to configure a static route for the NMS server, that is, configure the IP address of the next hop.

Step 2 Configure the default VLAN (native VLAN) for the ONT port.

Run the **ont port native-vlan** command to configure the default VLAN for the ONT port. By default, the default VLAN ID of the ONT port is 1.

- If the packets reported from a user (such a PC) to the ONT are untagged, the packets are tagged with the default VLAN of the port on the ONT and then reported to the OLT.
- If the packets reported from a user to the ONT are tagged, you need to configure the port VLAN of the ONT to be the same as the VLAN in the user tag. The packets are not tagged with the default VLAN of the port on the ONT but are reported to the OLT with the user tag.

Step 3 Bind an alarm profile.

Run the **ont alarm-profile** command bind an alarm profile. Ensure that Configuring a GPON ONT Alarm Profile is completed before the configuration.

Step 4 Activate the ONT.

Run the **ont activate** command to activate the ONT. The ONT can transmit services only when it is in the activated state.

After being added, the ONT is in the activated state by default. The step is required only when the ONT is in the deactivated state.

Step 5 Query the ONT status.

Run the **display ont info** command to query the ONT running status, configuration status, and matching status.

----End

Example

To add five ONTs in offline mode with password authentication mode (ONT passwords are 0100000001-0100000005), set the discovery mode of password authentication to always-on, and bind line profile 10 and service profile 10, do as follows:

```
huawei(config)#interface gpon 0/2
huawei(config-if-gpon-0/2)#ont add 0 password-auth 0100000001 always-on omci
ont-lineprofile-id 10 ont-srvprofile-id 10
huawei(config-if-gpon-0/2)#ont add 1 password-auth 0100000002 always-on omci
ont-lineprofile-id 10 ont-srvprofile-id 10
huawei(config-if-gpon-0/2)#ont add 2 password-auth 0100000003 always-on omci
ont-lineprofile-id 10 ont-srvprofile-id 10
huawei(config-if-gpon-0/2)#ont add 3 password-auth 0100000004 always-on omci
ont-lineprofile-id 10 ont-srvprofile-id 10
huawei(config-if-gpon-0/2)#ont add 4 password-auth 0100000005 always-on omci
ont-lineprofile-id 10 ont-srvprofile-id 10
```

To add an ONT that is managed by the OLT through the OMCI protocol, confirm this ONT according to the SN 3230313185885B41 automatically reported by the system, and bind the ONT with line profile 3 and service profile 3 that match the ONT, do as follows:

```
huawei(config)#interface gpon 0/2
huawei(config-if-gpon-0/2)#port 0 ont-auto-find enable
huawei(config-if-gpon-0/2)#ont confirm 0 sn-auth 3230313185885B41 omci
ont-lineprofile-id 3 ont-srvprofile-id 3
```

To add an ONU that is managed as an independent NE and whose SN is known as 3230313185885641, bind the ONU with line profile 4 that matches the ONU, configure the NMS parameters for the ONU, and set the management VLAN to 100, do as follows:

```
huawei(config)#snmp-profile add profile-id 1 v2c public private 10.10.5.53 161 huawei
huawei(config)#interface gpon 0/2
huawei(config-if-gpon-0/2)#ont add 0 2 sn-auth 3230313185885641 snmp
ont-lineprofile-id 4
huawei(config-if-gpon-0/2)#ont ipconfig 0 2 static ip-address 10.20.20.20 mask
255.255.255.0 gateway 10.10.20.1 vlan 100
huawei(config-if-gpon-0/2)#ont snmp-profile 0 2 profile-id 1
huawei(config-if-gpon-0/2)#ont snmp-route 0 2 ip-address 10.10.20.190 mask
255.255.255.0 next-hop 10.10.20.100
```

2.13.4 Configuring a GPON Port

To work normally and carry the service, a GPON port must be enabled first. This topic describes how to enable a GPON port and configure related attributes of the port.

Default Configuration

Table 2-17 lists the default settings of the GPON port.

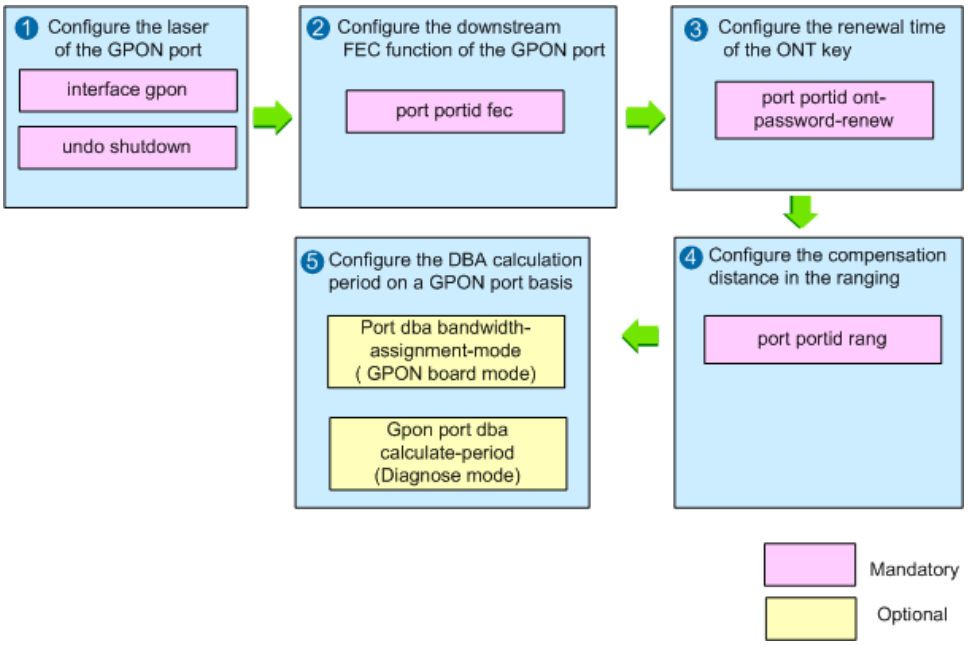
Table 2-17 Default settings of the GPON port

Parameter	Default Setting
GPON port	Enabled
Downstream FEC function of the GPON port	Disabled
Compensation distance range of the GPON port ranging	Minimum logical distance: 0 km; maximum logical distance: 20 km

Configuration Process

Figure 2-58 shows the process of configuring a GPON Port.

Figure 2-58 Process of configuring a GPON Port



Procedure

Run the **interface gpon** command to enter the GPON mode.

Step 1 Configure the laser of the GPON port.

- Run the **undo shutdown** command to enable the laser of the GPON port. By default, the laser of the GPON port is enabled and the GPON port is available. In this case, skip this step.
- If the GPON port is not to be used, run the **shutdown** command to disable the laser of the GPON port.

NOTICE

Disabling a PON port that carries services will cause the interruption of such services.

Step 2 Configure the downstream FEC function of the GPON port.

Run the **port *portid* fec** command to configure the FEC function of the GPON port. By default, the FEC function is disabled.

NOTE

- FEC is to insert redundant data into normal packets so that the line has certain error tolerance. Some bandwidth, however, must be consumed. Enabling FEC enhances the error correction capability of the line but at the same time occupies certain bandwidth. Determine whether to enable FEC according to the actual line planning.
- If a large number of ONTs are already online, enabling FEC on the GPON port may cause certain ONTs to go offline. Therefore, it is suggested that FEC should not be enabled on a GPON port that connects to online ONTs.

Step 3 Configure the renewal time of the ONT key.

Run the **port *portid* ont-password-renew** command to configure the interval for renewing the ONT key. To ensure the system security, the ONT key renewal must be configured.

Step 4 Configure the compensation distance in the ranging.

Run the **port *portid* range** command to configure the compensation distance range of the GPON port ranging. By default, the minimum logical distance is 0 km, and the maximum logical distance is 20 km. The difference between the minimum logical distance and the maximum logical distance must not exceed 20 km.

Step 5 (Optional) Configure the DBA calculation period on a GPON port basis.

When different GPON ports provide different access services, the bandwidth delays on these ports are different. In this case, the DBA calculation period needs to be configured on a GPON port basis.

1. In GPON board mode, run the **port dba bandwidth-assignment-mode** command to configure the DBA mode on a GPON port.
2. In diagnose mode, run the **gpon port dba calculate-period** command to configure the DBA calculation period on the GPON port.

NOTE

- The DBA calculation period on a GPON port can be configured only when the DBA mode is set to **manual** on this GPON port.
- By default, the DBA mode on a GPON port is **default**, which means the global DBA mode is used as the bandwidth assignment mode for the GPON port. In this case, if the global DBA mode is modified by running the **gpon dba bandwidth-assignment-mode** command, the bandwidth assignment mode on the GPON port is also modified. If the DBA mode on a GPON port is not **default**, the bandwidth assignment mode on the GPON port is not affected by the global DBA mode.
- If ONTs are configured on a GPON port, modifying the DBA mode is not allowed on this GPON port.
- For the TDM service, the DBA mode must be set to **min-loop-delay**.

----End

Example

Assume that the key renew interval of the ONT under the port is 10 hours, the minimum compensation distance of ranging is 10 km, and the maximum compensation distance of ranging is 15 km. To enable the FEC function of GPON port 0/2/0, do as follows:

```
huawei(config)#interface gpon 0/2
huawei(config-if-gpon-0/2)#port 0 fec enable
huawei(config-if-gpon-0/2)#port 0 ont-password-renew 10
huawei(config-if-gpon-0/2)#port 0 range min-distance 10 max-distance 15
  This command will result in the ONT's re-register in the port.
  Are you sure to execute this command? (y/n)[n]: y
```

To set the global DBA mode to **min-loop-delay**, DBA mode on GPON port 0/2/0 to **manual**, and DBA calculation period to **4**, do as follows:

```
huawei(config)#gpon dba bandwidth-assignment-mode min-loop-delay
huawei(config)#interface gpon 0/2
huawei(config-if-gpon-0/2)#port dba bandwidth-assignment-mode 0 manual
huawei(config-if-gpon-0/2)#quit
huawei(config)#diagnose
huawei(diagnose)%%gpon port dba calculate-period 0/2/0 4
```

2.13.5 Configuring GPON Type B Single-Homing Protection

This section describes how to configure GPON type B single-homing protection on an OLT to implement GPON port 1+1 redundancy backup, which ensures that services are not interrupted if a fault occurs on the OLT's PON port or backbone fiber.

Precautions

After GPON type B single-homing protection is configured, the service configurations on an optical network unit (ONU) remain unchanged and data is transmitted or received over the primary GPON port.

The GPON type B protection function is incompatible with the GPON type C protection functions. Only one of the two functions can be enabled on a network.

Procedure

Run the **protect-group** command to create a protection group for GPON access ports.

NOTE

1. Set **protect-target** to **gpon-uni-port**.
2. The working mode of the members in the protection group can only be **timedelay**.

Step 1 Run the **protect-group member** command to add a protection member to the protection group.

NOTE

- When adding a protection group member, add a working member and then a protection member.
- Protection group members can be added only based on ports.
- Protection group members can be GPON ports on different boards of the same type.

Step 2 Run the **protect-group enable** command to enable the protection group.

A created protection group is disabled by default.

Step 3 Run the **display protect-group** command to query the information about the protection group and all the members in the protection group.

NOTE

Bind a PPPoE single MAC address pool to a protection group if PPPoE single MAC is enabled. To do so, run the **bind mac-pool single-mac** command in the protect-group mode. Otherwise, the PPPoE service carried over the GPON port is interrupted when a protection switchover is performed. In this case, users need to dial numbers up again to go online and the service interruption time is based on BRAS configurations. This may fail to meet the switchover performance requirements of no longer than 50 ms for a protection switchover.

----End

Result

After the configuration, the primary GPON port on the OLT works in active mode and the secondary GPON port works in standby mode.

An automatic switching can be triggered by any of the following conditions:

- OLT GPON port failure
- Fractures of optical fibers

- Quality deterioration of lines

Example

The following configurations are used as an example to configure GPON type B single-homing protection on the OLT:

- Protection group members: 0/2/0 and 0/2/1 (on the same GPON service board)
- Primary port: 0/2/0
- Secondary port: 0/2/1

```
huawei(config)#protect-group 0 protect-target gpon-uni-port workmode timedelay
huawei(protect-group-0)#protect-group member port 0/2/0 role work
huawei(protect-group-0)#protect-group member port 0/2/1 role protect
huawei(protect-group-0)#protect-group enable
```

The following configurations are used as an example to configure type B single-homing protection on the OLT:

- Protection group members: 0/3/1 and 0/4/1 (on different GPON service boards)
- Primary port: 0/3/1
- Secondary port: 0/4/1

```
huawei(config)#protect-group 0 protect-target gpon-uni-port workmode timedelay
huawei(protect-group-0)#protect-group member port 0/3/1 role work
huawei(protect-group-0)#protect-group member port 0/4/1 role protect
huawei(protect-group-0)#protect-group enable
```

2.13.6 Configuring GPON Type B Dual-Homing Protection

Dual-homing GPON type B protection is enhanced based on single-homing GPON type B protection. Each of the two OLTs in a dual-homing protection group connects to a feeder fiber for remote disaster recovery.

Prerequisites

An ONU has been added to the active and standby OLTs by running the **ont add** command. All ONU profiles, such as DBA profile and line profiles, are the same on the active and standby OLTs.

Precautions

On a dual-homed network, two OLTs are in active/standby state, and they cannot forward packets at the same time. Users must manually configure the same service data on the two OLTs so that the ONU can rapidly switch services from the active OLT to the standby one, minimizing service interruption duration.

GPON type B protection is incompatible with GPON type C protection. Therefore, do not configure the two types of protection on the same ONU.

Procedure

Configure a dual-homing GPON type B protection group on the active OLT.

1. Run the **dual-parenting local-node** command to configure the local IP address, TCP port number, and key.
2. Run the **dual-parenting peer-node** command to configure the peer IP address, TCP port number, and key.
3. Run the **dual-parenting sync** command to enable dual-homing synchronization.
4. Run the **protect-group** command to create a protection group.
   - Set **protect-target** to **gpon-uni-port**.
   - Set the working mode of the protection group to **dual-parenting**.
5. Run the **protect-group member** command to add a member to the protection group.

   NOTE
   After a member is added to a dual-homing protection group, the group is automatically enabled.
6. Run the **peer-group-member** command to configure the peer member of the protection group.

Step 1 Configure a dual-homing GPON type B protection group on the standby OLT.

1. Run the **dual-parenting local-node** command to configure the local IP address, TCP port number, and key.
2. Run the **dual-parenting peer-node** command to configure the peer IP address, TCP port number, and key.
3. Run the **dual-parenting sync** command to enable dual-homing synchronization.
4. Run the **protect-group** command to create a protection group.
   - Make sure that the description of the protection groups created on the active and standby OLTs are the same. Run the **description** command to configure the description of a PG.
   - Set **protect-target** to **gpon-uni-port**.
   - Set the working mode of the protection group to **dual-parenting**.
5. Run the **protect-group member** command to add a member to the protection group.
6. Run the **peer-group-member** command to configure the peer member of the protection group.
7. Run the **protect-group enable** command to enable the protection group.

Step 2 (Optional) Run the **uplink-monitor port** command to associate the protection groups with the uplink Ethernet port status.

- If an Ethernet link aggregation group has been configured, make sure that the Ethernet port associated with the dual-homing protection groups is the master port in the aggregation group.
- Run the **uplink-monitor bfd** command to associate the protection groups with a BFD session. For instructions about how to configure a BFD session, see **Configuring a BFD Session**.
- Run the **uplink-monitor mep** command to associate the protection groups with an MEP session. For instructions about how to configure an MEP session, see **Configuring CFM**.

----End

Result

After the configuration, both active and standby OLTs are in working state, and they both check link status for determining a protection switchover.

An automatic switchover can be triggered by any of the following conditions:

- Optical fiber cut from the active OLT
- Active OLT PON port fault
- Active OLT fault
- Active line quality deterioration
- Active OLT uplink failure

Example

The following section uses an example to describe how to configure a dual-homing GPON type B protection group on the active OLT (huawei_A) and standby OLT (huawei_B), respectively:

- GPON service port on the two OLTs: 0/2/1
- Index of the protection groups on the two OLTs: 1
- Associated port with the dual-homing protection groups: 0/19/1
- IP address of the local huawei_A and peer huawei_B: 192.168.68.1; TCP port number: 6076; key: work_4234
- IP address of the local huawei_B and peer huawei_A: 192.168.68.8; TCP port number: 6076; key: protect_4234

```
Configuration on huawei_A:
huawei_A(config)#dual-parenting local-node ip-address 192.168.68.1 port 6076 key work_4234
huawei_A(config)#dual-parenting peer-node standby ip-address 192.168.68.8 port 6076 key protect_4234
huawei_A(config)#dual-parenting sync enable
huawei_A(config)#protect-group 1 protect-target gpon-uni-port workmode dual-parenting
huawei_A(protect-group-1)#protect-group member port 0/2/1 role work
huawei_A(protect-group-1)#peer-group-member peer-node standby peer-port 0/2/1
huawei_A(protect-group-1)#uplink-monitor port 0/19/1
huawei_A(protect-group-1)#quit
Configuration on huawei_B:
huawei_B(config)#dual-parenting local-node ip-address 192.168.68.8 port 6076 key protect_4234
huawei_B(config)#dual-parenting peer-node active ip-address 192.168.68.1 port 6076 key work_4234
huawei_B(config)#dual-parenting sync enable
huawei_B(config)#protect-group 1 protect-target gpon-uni-port workmode dual-parenting
huawei_B(protect-group-1)#protect-group member port 0/2/1 role protect
huawei_B(protect-group-1)#peer-group-member peer-node active peer-port 0/2/1
huawei_B(protect-group-1)#protect-group enable
huawei_B(protect-group-1)#uplink-monitor port 0/19/1
huawei_B(protect-group-1)#quit
```

2.13.7 Configuring GPON Type C Single-Homing Protection

This section describes how to configure GPON type C single-homing protection. Each optical network unit (ONU) provides to GPON uplink ports and connects to two GPON ports on an OLT through different optical splitters. This protects feeder and drop fibers and ensures high network reliability.

Context

After GPON type C single-homing protection is configured, the service configurations on the ONUs remain unchanged and data is transmitted or received over the primary uplink ports on the ONUs and the primary GPON port on the OLT.

Precautions

The GPON type C single-homing protection function is incompatible with the GPON type B protection and GPON type C dual-homing protection functions. Only one of the three functions can be enabled on a network.

Procedure

Run the **ont add** command to add a work-side ONU.

Step 1 Run the **ont add** *portid ontid* **protect-side** command to add a protect-side ONU.

Ensure that the **protect-side** parameter is selected.

Step 2 Run the **protect-group protect-target gpon-uni-ont** command to add a protection group. The working mode of the protection group can only be **portstate**.

Step 3 Run the **protect-group member** command to add the work-side ONU to the protection group as a working member.

Step 4 Run the **protect-group member** command to add the protect-side ONU to the protection group as a protection member.

NOTE
Ensure that **ont** *ontid* value is the ONT ID specified in Step 1.

Step 5 Run the **protect-group enable** command to enable the protection group.

A created protection group is disabled by default.

----End

Result

The OLT will switch the services on the primary GPON port to the secondary GPON port if one of the following requirements is met:

- Loss of signal (LOS) occurs in the input direction.
- The OLT or ONU hardware is faulty.

Example

The following configurations are used as an example to configure GPON type C single-homing protection on an OLT and ONUs:

- Ports on the same GPON service board: 0/2/0 and 0/2/1
- Work-side link: connected to port 0/2/0
- Protect-side link: connected to port 0/2/1
- ONU ID: 0
- ONU authentication mode: SN; SN: hwhw-10101500; management mode: SNMP
- ID of the line profile bound to the ONU: 10

```
huawei(config)#interface gpon 0/2
huawei(config-if-gpon-0/2)#ont add 0 0 sn-auth hwhw-10101500 snmp ont-lineprofile-id
10
huawei(config-if-gpon-0/2)#ont add 1 0 protect-side
huawei(config-if-gpon-0/2)#quit
huawei(config)#protect-group protect-target gpon-uni-ont workmode portstate
huawei(protect-group-1)#protect-group member port 0/2/0 ont 0 role work
huawei(protect-group-1)#protect-group member port 0/2/1 ont 0 role protect
huawei(protect-group-1)#protect-group enable
huawei(protect-group-1)#quit
```

2.13.8 Configuring GPON Type C Dual-Homing Protection

This section describes how to configure type C dual-homing protection. Type C dual-homing protection is enhanced based on type C single-homing protection. Type C dual-homing protection protects any node between an optical network unit (ONU) and the two dual-homed OLTs.

Precautions

Compared with GPON type C single-homing protection, GPON type C dual-homing protection features enhanced protection capabilities but more complicated networking and higher deployment costs. For configuration examples of the GPON type C dual-homing protection, see the *FTTx Solution Configuration Guide*.

The GPON type C dual-homing protection function is incompatible with the GPON type B protection and GPON type C single-homing protection functions. Only one of the three functions can be enabled on a network.

Procedure

Configure GPON type C dual-homing protection on the primary OLT.

1. Run the **ont add** command to add a work-side ONU.
2. Run the **protect-group** command to create a protection group.
   - Set **protect-target** to **gpon-uni-ont**.
   - The working mode of the protection group can only be **dual-parenting**.
3. Run the **protect-group member** command to add a working member to the protection group.

NOTE
After this step is performed, the protection group is automatically enabled.

Step 1 Configure type C dual-homing protection on the secondary OLT.

1. Run the **ont add** command to add a protect-side ONU.

NOTE
Ensure that all profiles on the ONUs connected to the primary and secondary OLTs are the same. The profiles include the dynamic bandwidth allocation (DBA) profile and line profile.

2. Run the **protect-group** command to create a protection group.
   - Make sure that the description of the protection groups created on the active and standby OLTs are the same. Run the **description** command to configure the description of a PG.
   - Set **protect-target** to **gpon-uni-ont**.
   - The working mode of the protection group can only be **dual-parenting**.
3. Run the **protect-group member** command to add a protection member to the protection group.
4. Run the **protect-group enable** command to enable the protection group.

----End

Result

After the configuration, both the primary and secondary OLTs work in active mode. The ONU checks the status of the links to the primary and secondary OLTs.

The OLT will switch the services carried over the primary link to the secondary link if one of the following requirements is met:

- Loss of signal (LOS) occurs in the input direction.
- The OLT or ONU hardware is faulty.

Example

The following configurations are used as an example to configure GPON type C dual-homing protection on two OLTs:

- Primary OLT: huawei_A; secondary OLT: huawei_B
- GPON service ports: 0/2/1 on the two OLTs
- ID of the protection groups: 1
- ONU ID: 0
- ONU authentication mode: SN; SN: hwhw-10101500; management mode: SNMP
- ID of the line profile bound to the ONU: 10

```
huawei_A configurations:
huawei_A(config)#interface gpon 0/2
huawei_A(config-if-gpon-0/2)#ont add 1 0 sn-auth hwhw-10101500 snmp ont-lineprofile-id
10
huawei_A(config-if-gpon-0/2)#quit
huawei_A(config)#protect-group 1 protect-target gpon-uni-ont workmode dual-parenting
huawei_A(protect-group-1)#protect-group member port 0/2/1 ont 0 role work
huawei_B configurations:
huawei_B(config)#interface gpon 0/2
```

```
huawei_B(config-if-gpon-0/2)#ont add 1 0 sn-auth hwhw-10101500 snmp ont-lineprofile-id
10
huawei_B(config-if-gpon-0/2)#quit
huawei_B(config)#protect-group 1 protect-target gpon-uni-ont workmode dual-parenting
huawei_B(protect-group-1)#protect-group member port 0/2/1 ont 0 role protect
huawei_B(protect-group-1)#protect-group enable
```

2.14 Reference Standards and Protocols

The reference standards and protocols of the GPON feature are as follows:

Standard No.	Description
ITU-T G.984.1	General Characteristics. This protocol mainly describes the basic features and major protection modes of GPON.
ITU-T G.984.2	Physical Media Dependent (PMD) Layer Specification. This protocol mainly describes the PMD layer parameters, including physical parameters (such as the transmit optical power, receiver sensitivity, and overload optical power) of optical transceivers, and also defines optical budget of different levels, for example, the most common Class B+.
ITU-T G.984.3	Transmission Convergence Layer Specification. This protocol mainly describes the TC layer specifications, including the upstream and downstream frame structures and GPON principle.
ITU-T G.984.4	ONT Management And Control Interface Specification. This protocol mainly describes the GPON management and maintenance protocols, such as OAM, PLOAM, and OMCI.
ITU-T G.984.5	Enhancement Band. This protocol mainly describes the GPON wavelength planning, including reserving bands for next-generation PON.
ITU-T G.984.6	Reach Extension. This protocol mainly describes several long reach PON schemes for extending GPON transmission distance.
ITU-T G.988	ONU management and control interface (OMCI) specification.
TR-156	Using GPON Access in the context of TR-101.

3 10G GPON

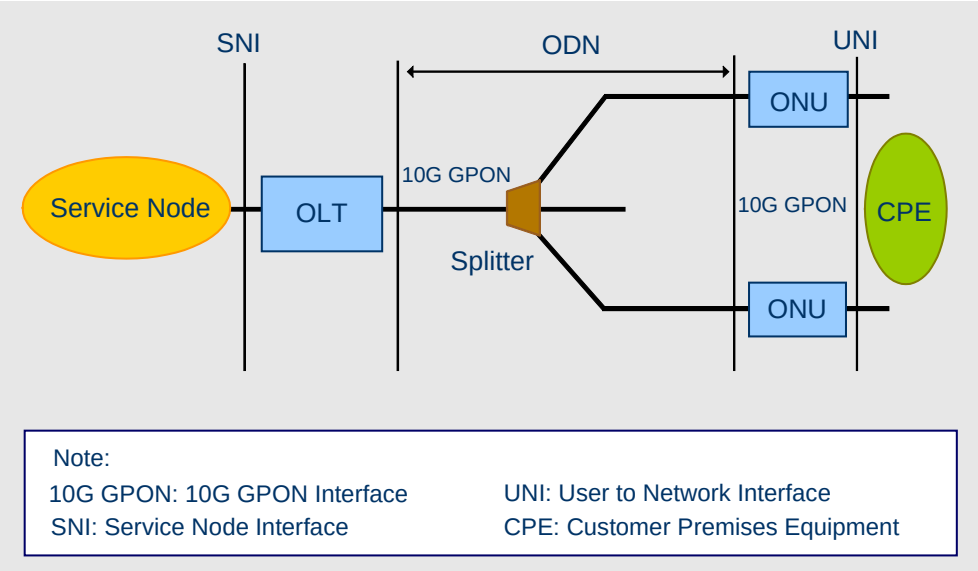
About This Chapter

3.1 Overview

Networking Diagram

A 10G GPON network is of the point-to-multipoint (P2MP) type, which is the same as that of a GPON network. Figure 3-1 shows a 10G GPON networking diagram.

Figure 3-1 Networking Diagram



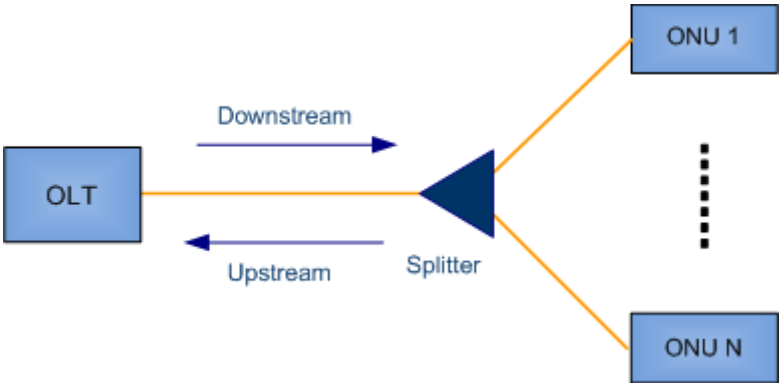
The 10G GPON network contains an optical line terminal (OLT), optical network units (ONUs), and an optical distribution network (ODN).

- The Optical line terminal (OLT) is an aggregation device located at the central office (CO) for terminating the PON protocol.
- Optical network units (ONUs) are located on the user side, providing various types of ports for connecting to user terminals. The OLT and ONUs are connected through a passive ODN for communication.
- The Optical distribution network (ODN) is composed of passive optical components (POS) such as optical fibers, and one or more passive optical splitters. The ODN provides optical channels between the OLT and ONUs. It interconnects the OLT and ONUs and is highly reliable.

Transmit Principles

10G GPON uses wavelength division multiplexing (WDM) to transmit data in different wavelengths on an ODN network. Figure 3-2 shows the working principles.

Figure 3-2 Transmit Principles



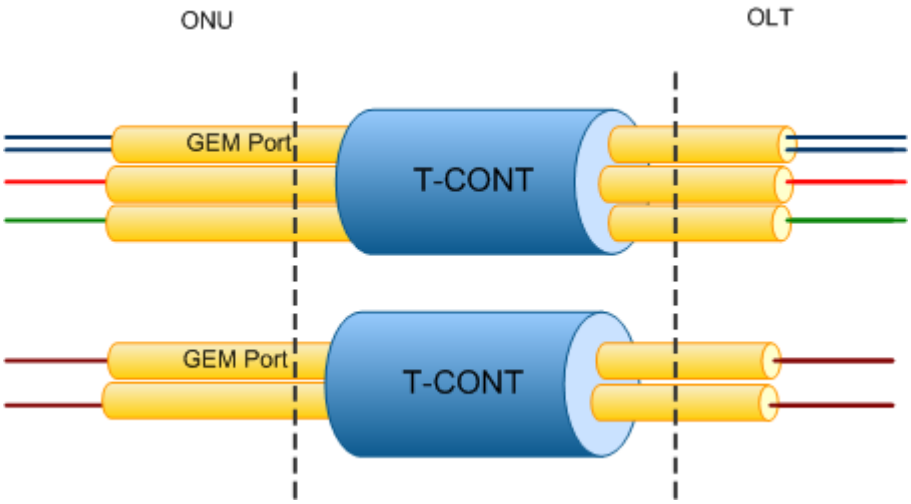
- Data is broadcast in the downstream direction.
- Data is transmitted in the TDMA mode (based on timeslots) in the upstream direction.

3.2 Basic Concepts

Service Multiplexing

GEM ports and T-CONTs divide a PON network into virtual connections for service multiplexing, as shown in Figure 3-3.

Figure 3-3 Working principles of service multiplexing in an 10G GPON system



GEM Port

A GPON encapsulation mode (GEM) port is a virtual service channel that carries a service flow between the OLT and an ONU in an 10G GPON system. The GEM port is similar to the virtual connection (identified by VPI/VCI) in asynchronous transfer mode (ATM). VPI is the acronym for virtual path identifier and VCI is the acronym for virtual channel identifier.

- Each GEM port is identified by a unique XGEM port ID.
- The XGEM port ID is globally allocated according to the 10G GPON port by the OLT.
- A GEM port can carry one or more types of services.

T-CONT

A transmission container (T-CONT) is the basic control unit of upstream service flows in an 10G GPON system, and is also the unit for carrying service flows in the upstream direction. All the GEM ports are mapped to T-CONTs, and the OLT uses dynamic bandwidth allocation (DBA) to schedule upstream transmission.

- A T-CONT can carry one or more GEM ports according to user configurations.
- A T-CONT is identified uniquely by Alloc-ID.
- The Alloc-ID is allocated according to the 10G GPON port by the OLT.
- An ONU supports multiple T-CONTs configured for various service types.

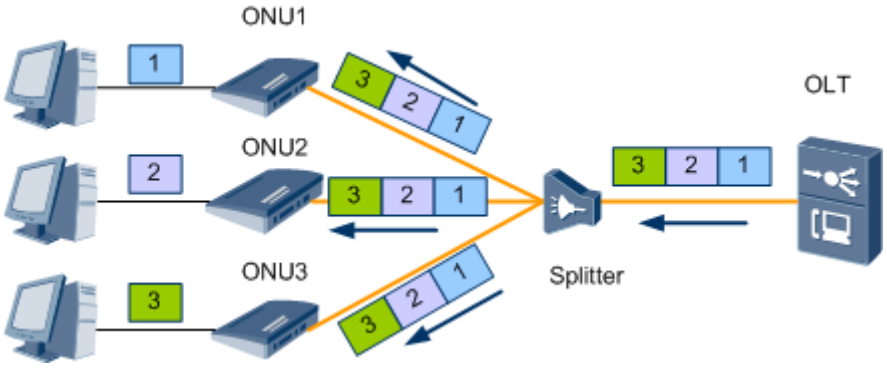
3.3 Working Principle

3.3.1 Working Principles of Downstream

Working Principles for Downstream Transmission

Figure 3-4 shows the 10G GPON working principles for downstream transmission.

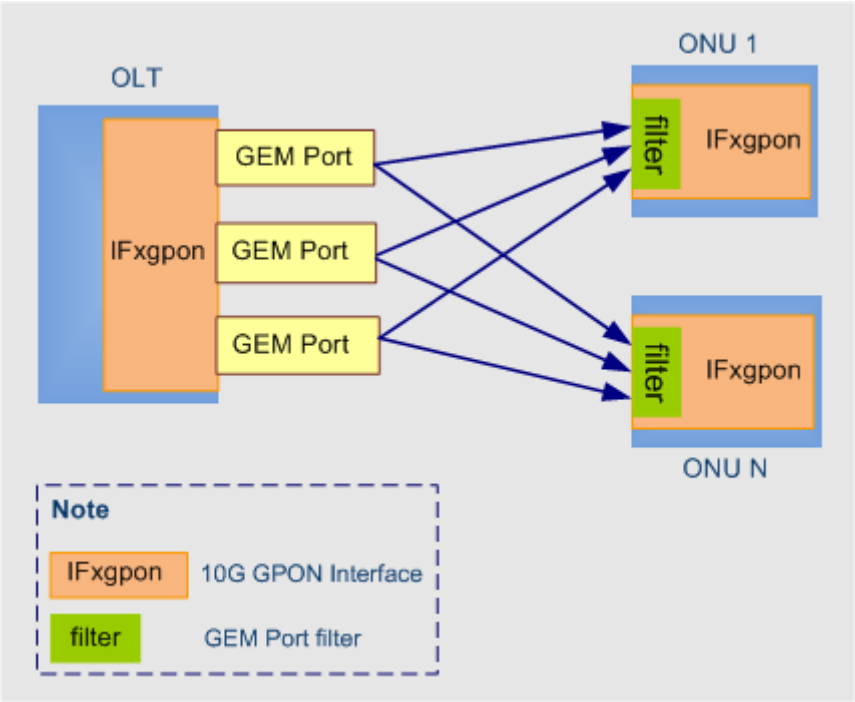
Figure 3-4 Working principles for downstream transmission



In the downstream direction, the OLT broadcasts data to all ONUs and the ONUs receive only desired data.

Data flow forwarding in the downstream direction

Figure 3-5 Data flow forwarding in the downstream direction



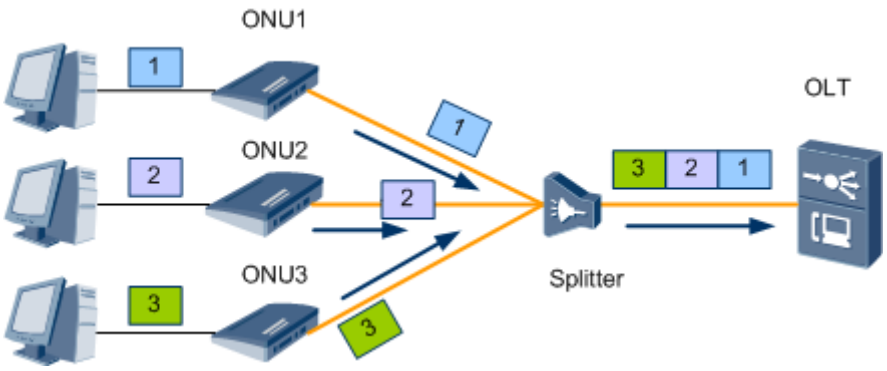
- On the OLT, data flows are encapsulated into GEM ports in service processing units.
- The OLT broadcasts the data to in the GEM ports to all ONUs.
- The ONU determines whether to process or discard the data according to the XGEM port ID.

3.3.2 Working Principle of Upstream

Working Principles for Upstream Transmission

Figure 3-6 shows the 10G GPON working principles for upstream transmission.

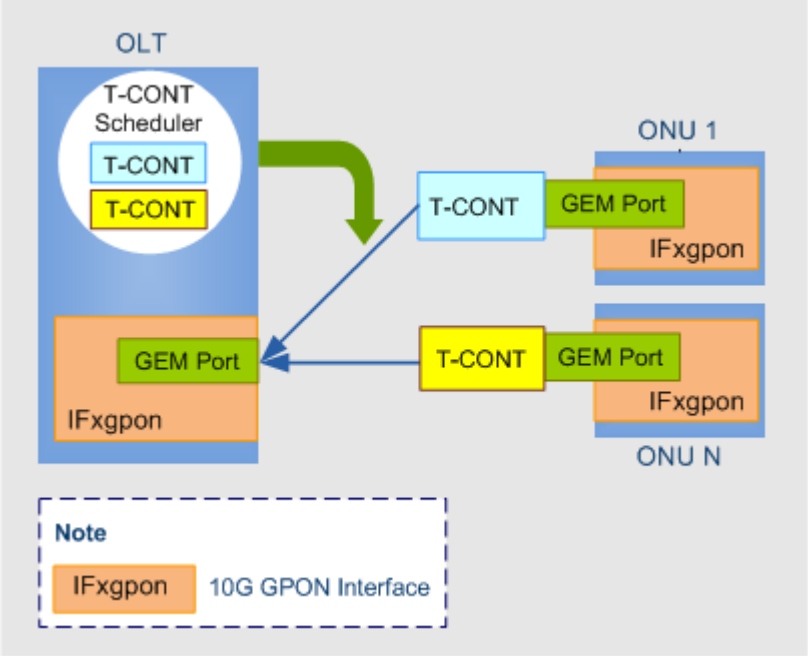
Figure 3-6 Working principles for upstream transmission



In the upstream direction, an ONU sends data to the OLT using an allocated timeslot. Such transmission ensures that all ONUs send data in a permitted sequence, which prevents upstream data collision.

Data flow forwarding in the upstream direction

Figure 3-7 Data flow forwarding in the upstream direction



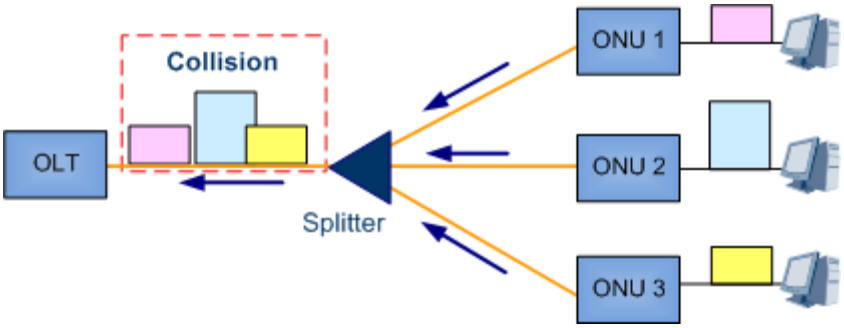
- On the ONU, data flows are encapsulated into GEM ports and mapped to transmission containers (T-CONTs).
- The ONU sends data flows to the OLT according to T-CONTs.
- The OLT decapsulates the data flows and sends them to service processing modules.

3.4 Key Technologies

3.4.1 Ranging

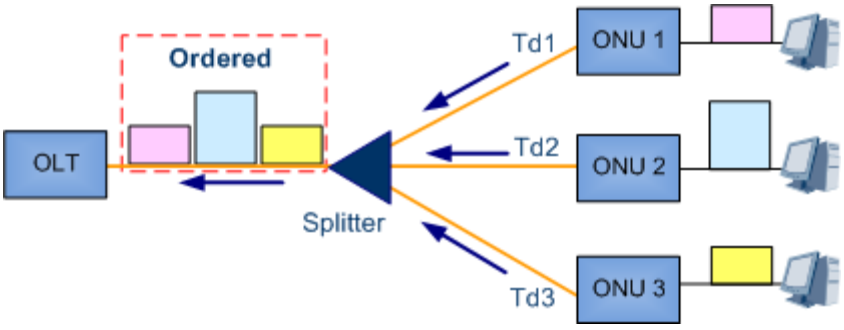
The logic reaches from optical network units (ONUs) to an optical line terminal (OLT) vary. The round trip delays (RTDs) between an OLT and ONUs also vary depending on time and environment. Therefore, collisions may occur when ONU sends data in TDMA mode (in this mode, only one of the ONUs connecting to a PON port sends data at a moment), as shown in Figure 3-8.

Figure 3-8 Cell transmission without ranging



To prevent the collisions, ranging is enabled when an ONU registers for the first time. The OLT measures the RTD of each ONU in the ranging process and calculates the equalization delay (EqD) of each ONU to ensure that the values of Teqd, which is equal to RTD plus EqD, of all ONUs connected to the same PON port are the same. Therefore, the logic reaches from ONUs to an OLT are the same, preventing collisions during upstream transmission.

Figure 3-9 Cell transmission with ranging



NOTE

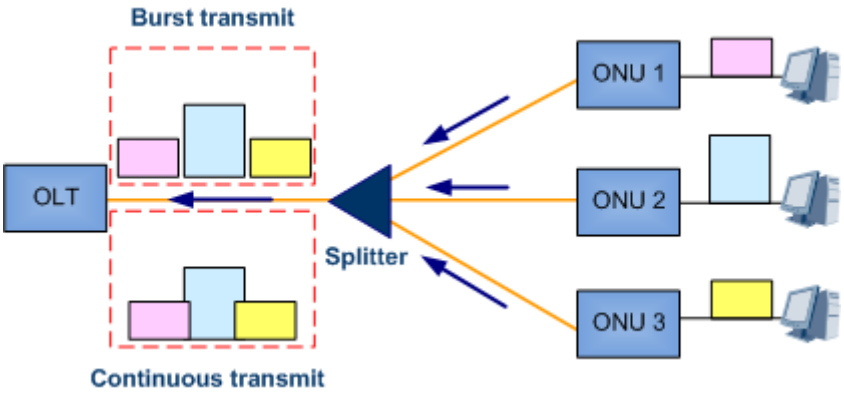
In the ranging process, the OLT must open a window and pause upstream transmission channels of other ONUs.

3.4.2 Burst Optical/Electrical Technology

In 10G GPON upstream direction, Time Division Multiple Access (TDMA) is used. An optical network unit (ONU) transmits data only within the allocated timeslots. In the timeslots that are not allocated to it, the ONU disables the transmission of its optical transceiver to prevent other ONUs from being affected. The optical line terminal (OLT) then receives the upstream data from each ONU in a burst manner based on timeslots. Therefore, to ensure normal running of the 10G GPON system.

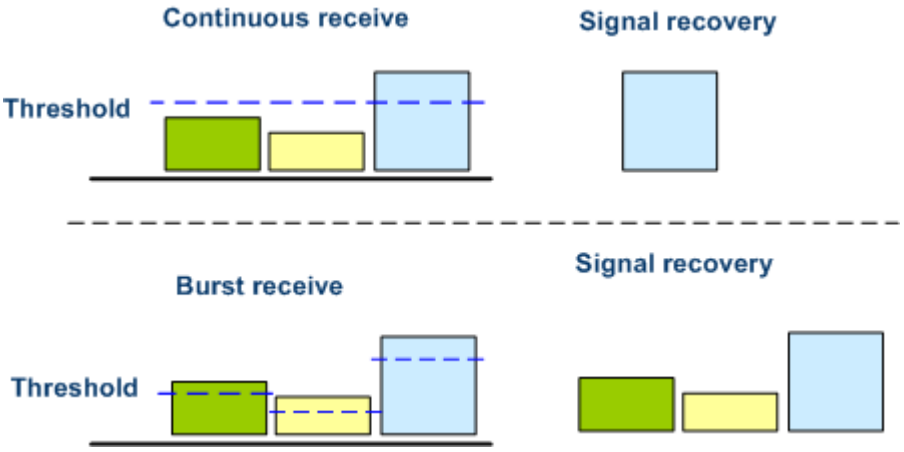
- Figure 3-10 shows the burst transmit function supported by ONU-side optical modules.
- Figure 3-11 shows the burst receive function supported by OLT-side optical modules.

Figure 3-10 Burst transmit function supported by ONU-side optical modules



Ranging can be implemented to prevent cells transmitted by different ONUs from conflicting with each other on the OLT. However, the ranging accuracy is ± 1 bit and the cells transmitted by different ONUs have a protection time of several bits (not a multiple of 1 bit). If the ONU-side optical modules do not support the burst transmit function, the transmitted signals overlap and distortion occurs.

Figure 3-11 Burst receive function supported by OLT-side optical modules



- The distance from each ONU to the OLT varies and therefore the optical signal attenuation varies for each ONU. As a result, the power and level of packets received by an OLT at different timeslots various.
- If the OLT-side optical modules do not support the burst receive function, the OLT may restore incorrect signals because only the level greater than the threshold is considered valid and the signals with the level lower than the threshold cannot be restored.

NOTE

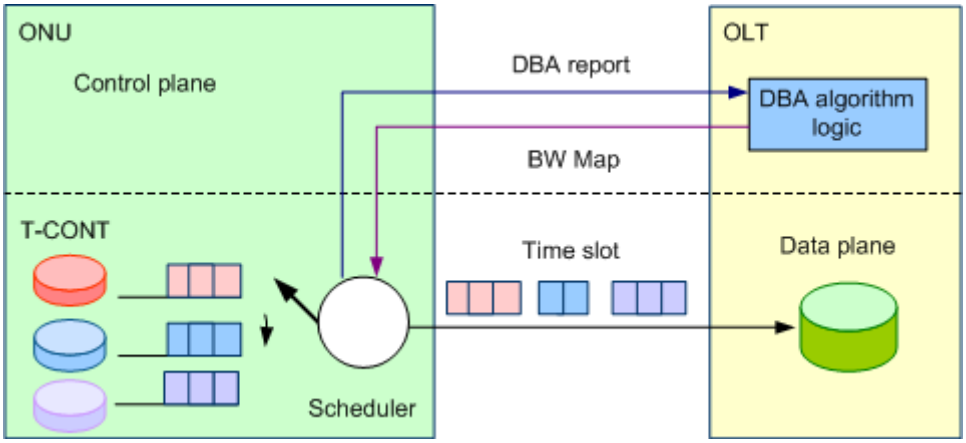
In the 10G GPON system, all data is broadcast downstream to ONUs. The transmission requires OLT-side optical modules to transmit optical signals continuously and ONU-side optical modules to receive optical signals continuously. Therefore, these optical modules are not required to support the burst receive and transmit function.

3.4.3 DBA

The OLT uses DBA to dynamically adjust the upstream bandwidth allocated to different ONUs to address the burst traffic on the ONUs, meeting the ONU upstream bandwidth requirements and improving the utilization of the PON upstream bandwidth.

Figure 3-12 shows the principles of DBA.

Figure 3-12 Principles of DBA

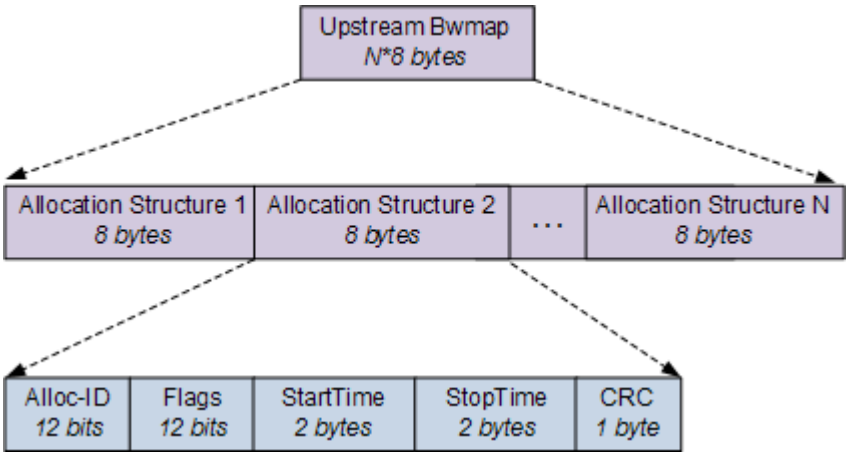


In the preceding figure,

- The DBA module in the OLT consistently collects DBA reports and uses the DBA algorithm to calculate the upstream bandwidth allocated to each ONU.
- The OLT sends the calculated result to each ONU using a bandwidth (BW) map.
- Each ONU transmits burst upstream data using permitted timeslots defined in the BW map.

A BW map allows an ONU to send upstream data. Figure 3-13 shows a BW map structure.

Figure 3-13 BW map structure



Highlights and Applications

- Based on ONUs' burst upstream service traffic, the OLT dynamically allocates an upstream bandwidth to each ONU in real time, improving upstream bandwidth utilization on PON ports.
- More users are supported on a PON port.
- Higher service bandwidths with burst requirements are supported than those before DBA is applied.

3.4.4 FEC

Context

Forward error correction (FEC) is mainly used for improving transmission quality of a line.

No ideal digital channel is available in practice. As a result, bit errors and jitter occur when digital signals are being transmitted over any transmission medium, deteriorating transmission quality on lines.

To resolve the problem, error correction mechanism is introduced.

- The mechanism can check and correct errors after data is transmitted to the peer end. such as FEC.
- The mechanism can check errors after data is transmitted to the peer end but not correct errors.

Highlight and Application

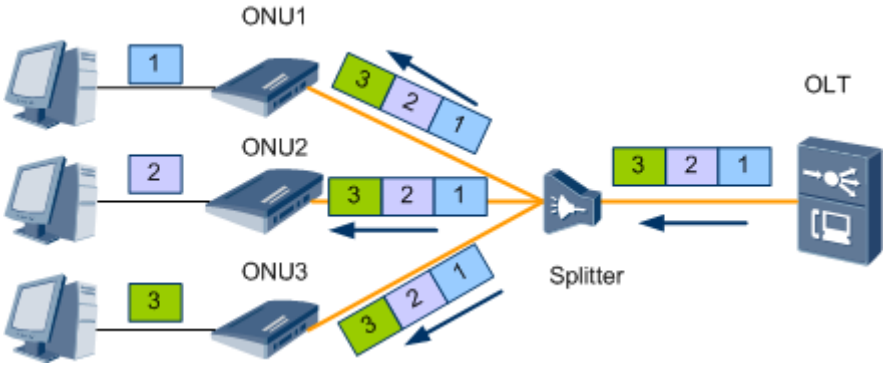
- Does not require retransmission and provides a high real-time performance
- Requires an additional bandwidth (Users must balance the transmission quality and bandwidth.)
- Checks and corrects errors after data is transmitted to the peer end, but does not apply to services for which retransmission is enabled
- Applies to data transmission on the network that has a poor quality
- Applies to services that have a low requirement on delay (The delay is large if retransmission is configured for services.)

3.4.5 Line Encryption

Context

In the PON system, downstream data is broadcast to all ONUs. As a result, downstream data destined for certain ONUs or all ONUs may be intercepted by illegal users.

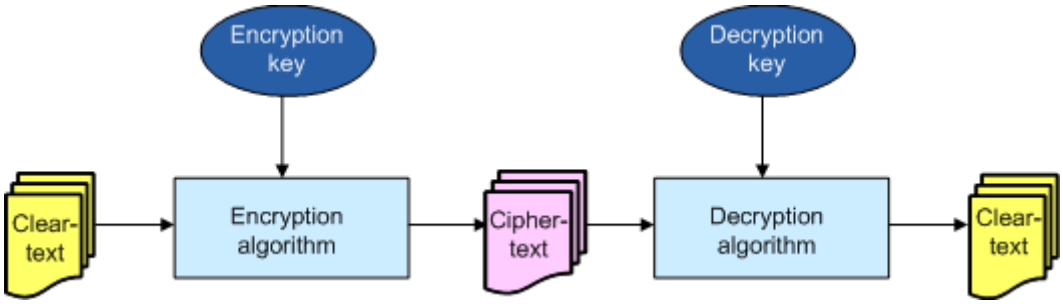
Figure 3-14 Downstream data transmit process



Working Principle

Line encryption technologies are required to eliminate the data theft risk, Figure 3-15 shows line encryption process.

Figure 3-15 Line encryption process



- The encryption algorithm to be used is the advanced encryption standard (AES).
- The 10G GPON systems use the AES-128 encryption algorithm.

Highlight and Application

- The line encryption algorithms neither increase overhead nor decrease bandwidth usages.
- The line encryption algorithms will not prolong transmission delays.
- Enable line encryption if the usage scenarios promote high security requirements.

3.5 Network Planning

Background Information

This section describes two common networking scenarios for evolving GPON to 10G GPON. You can select either of them based on the actual networking and service requirements.

Networking Scenario I — Pure 10G GPON Network

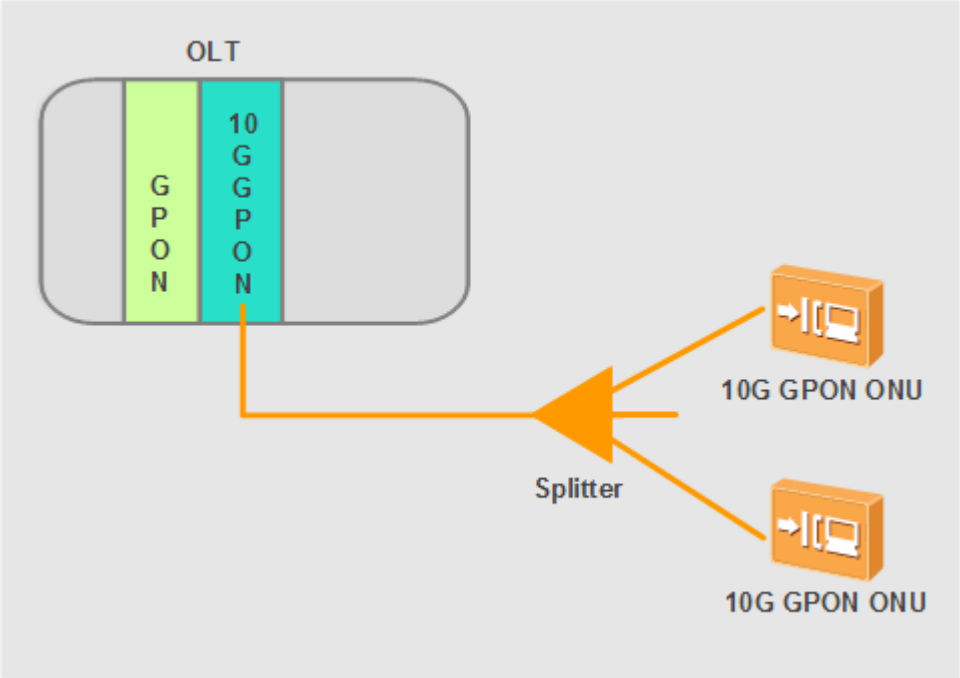
Description

A pure 10G GPON network contains only 10G GPON devices, including 10G GPON service board, 10G GPON optical network units (ONUs), and an optical distribution network (ODN). This scenario applies to a new 10G GPON FTTB or FTTC network. FTTB is the acronym for fiber to the building and FTTC is the acronym for fiber to the curb.

Network Diagram

Figure 3-16 shows a pure 10G GPON network.

Figure 3-16 Pure 10G GPON network



Characteristics

- Advantage: Only one type of network element (NE) (10G GPON devices) is involved, and the maintenance is easy.
- Disadvantage: A new ODN is required.

Networking Scenario II — Hybrid 10G GPON and GPON Network

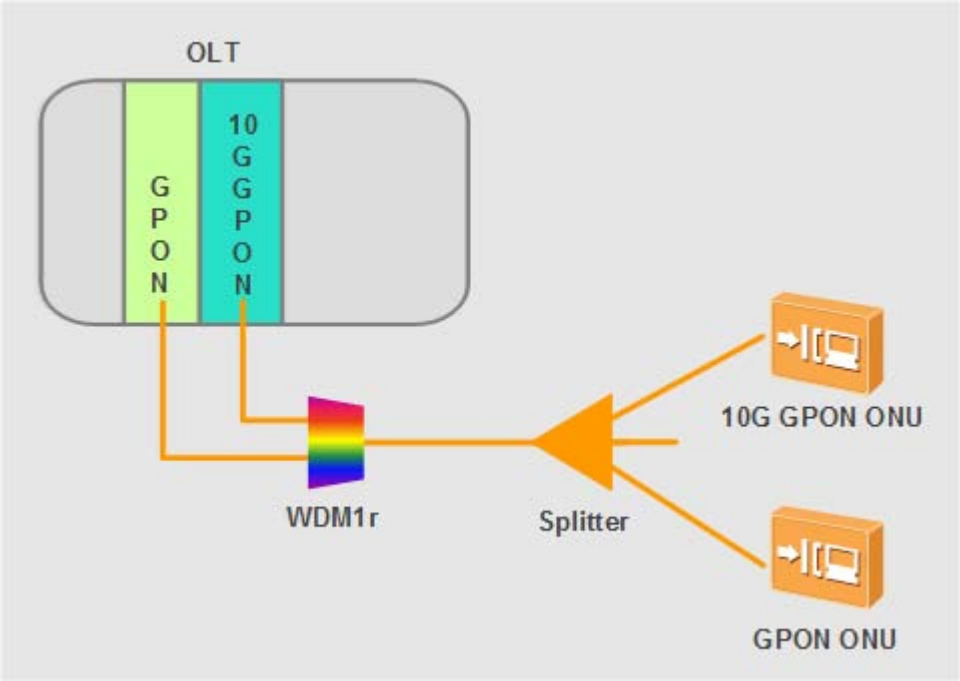
Description

A hybrid 10G GPON and GPON network contains 10G GPON and GPON NEs. These 10G GPON and GPON NEs share an ODN.

Network Diagram

Figure 3-17 shows a hybrid 10G GPON and GPON network.

Figure 3-17 Hybrid 10G GPON and GPON network



On a hybrid 10G GPON and GPON network, an OLT uses 10G GPON and GPON boards to receive services. 10G GPON and GPON NEs share an ODN but use different service wavelengths. Therefore, a passive wavelength multiplexing device (WDM1r) is required.

Characteristics

- Advantage: A GPON network is smoothly migrated to a 10G GPON network and the two networks share an ODN.
- Disadvantages:
  - A WDM1r device is required on the ODN to multiplex wavelengths. This operation requires reconstruction for existing ODN networks and optical fiber connections, which interrupts services.
  - Various types of NEs (10G GPON and GPON devices) are involved, and the maintenance is complicated.

NOTE
A hybrid network is complicated and therefore is not recommended.

Notes

10G GPON ONUs are compatible with GPON ONUs. During the usage, pay attention to the following points:

- On OLTs, 10G GPON access board support only 10G GPON ONUs.
- On OLTs, GPON access board support only GPON ONUs.

3.6 Configuration Guide

3.6.1 Configuring a Service Board

This section describes how to configure a 10G GPON service board.

Adding a Board

You can add a 10G GPON service board using either of the following methods:

- Manually inserting a board: When 10G GPON boards have been configured in the subrack, manually insert a required board.
- Adding a board offline: When a 10G GPON service board needs to be pre-configured, add a 10G GPON service board offline.

NOTE

- In global config mode, run the **board add** *frameid/slotidboard-type* command to add a 10G GPON service board.
- After a board is successfully added offline, the board status is **Failed**. However, you can still configure or query data on the board.
- After a user manually inserts a board (the board type must be the same as that of the board added offline), the board status changes to **Normal**. Data configured for the board takes effect immediately after the configuration.

Configuring the Working Mode of a Board

Configure the working mode of a 10G GPON service board based on actual requirements. The PON system supports the following modes:

- GPON mode
- XG-PON mode, that is, 10G GPON mode

Commands related to the working mode of a 10G GPON service board are as follows:

- In diagnosis mode, run the **display gpon board workmode** command to query the working mode of a 10G GPON service board.
- In diagnosis mode, run the **gpon board workmode** command to configure the working mode of a 10G GPON service board.

NOTE

A 10G GPON service board works in XG-PON mode by default.

3.6.2 Configuring Port Attributes

Automatic Discovery of an ONT

The 10G GPON system adds an ONT using either of the following methods:

- Adding an ONT offline: Before installing an ONT, manually add an ONT to the OLT and configure the ONT. After the ONT goes online, the OLT authenticates the ONT and issues configurations to the ONT.
- Adding an ONT online: After an ONT is installed, the OLT discovers the online ONT and adds and configures it.

The automatic discovery of an ONT applies when an ONT is added online. In this scenario, the installation time of an ONT is uncertain because the OLT periodically searches for online ONTs.

- The automatic discovery of an ONT connecting to 10G GPON ports is disabled by default.
- In 10G GPON interface mode, run the **port ont-auto-find** command to enable the automatic discovery of an ONT.
- In 10G GPON interface mode, run the **display ont autofind** command to query the automatic discovery of an ONT.

Laser

Run a command to enable or disable a laser for a 10G GPON port.

- The laser for a 10G GPON port is enabled by default.
- In 10G GPON interface mode, run the **shutdown** command to disable a laser.
- In 10G GPON interface mode, run the **undo shutdown** command to enable a laser.

NOTE

After a laser is disabled, all services carried on the port with the laser are interrupted. Exercise caution when disabling a laser.

3.7 Reference Standards and Protocols

The following lists XG-PON-related standards:

- ITU-T G.987.1: defines overall requirements in the next generation PON system.
- ITU-T G.987.2: provides physical medium dependent (PMD) specifications at the physical layer in the next generation PON system
- ITU-T G.987.3: provides GPON transmission convergence (GTC) specifications at the convergence layer in the next generation PON system. GPON is the acronym for gigabit-capable passive optical network.
- ITU-T G.988: provides optical network terminal management and control interface (OMCI) specifications in the next generation PON system

4 P2P Optical Access

About This Chapter

Point-to-point (P2P) Ethernet optical access refers to the P2P FTTX access provided by the P2P Ethernet optical access board and the ONT, which meets the requirements for the application of the next generation access device under the integration of video, voice, and data services.

4.1 P2P FE Optical Access

4.1.1 Introduction

Definition

Point-to-point (P2P) FE optical access means the point-to-point FTTH access provided by the MA5600T/MA5603T/MA5608T based on the combination between its P2P FE optical access board and the ONTs.

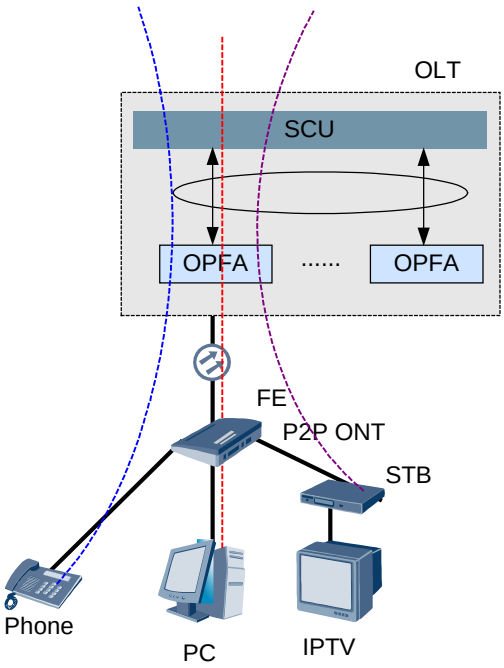
Purpose

P2P FE optical access solution provides P2P FTTH access services. It is especially suitable for the residential neighborhoods with fiber to the home, and can provide the bandwidth of 100 Mbit/s to satisfy the users' requirements for the next generation access equipment which integrates video, voice, and data services.

4.1.2 Principle

Figure 4-1 shows the implementation of the P2P FE optical access.

Figure 4-1 Implementation of P2P FE optical access



The upstream packets sent from the user end are processed as follows:

1. After modulation on the ONT, the upstream packets are sent to the P2P board of the MA5600T/MA5603T/MA5608T through a fiber.
2. The P2P board processes the upstream packets according to the user's configuration, and then sends the processed packets to the control board of the MA5600T/MA5603T/MA5608T through the backplane bus.
3. After receiving the packets, the control board forwards the packets to the upper layer network through the upstream port.

The downstream packets sent from the network end are processed as follows:

1. The downstream packets from the upper layer network reach the control board of the MA5600T/MA5603T/MA5608T through the upstream port.
2. The control board forwards the packets to the P2P interface board through the backplane bus according to the learning results during the upstream forwarding.
3. The P2P board processes the downstream packets, and sends the processed packets to the user end.

4.1.3 Reference Standards and Protocols

For the standards compliance of the P2P FE optical access feature, see "Standards Compliance" in the *MA5600T/MA5603T/MA5608T* Product Description.

4.2 GE P2P Optical Access

4.2.1 Introduction

Definition

GE point-to-point (P2P) Ethernet optical access is a mode in which P2P Ethernet optical access boards provide GE ports and coordinate with downstream devices to implement various optical access solutions for users. The solutions include FTTC/FTTB, FTTH, FTTO, and FTTM.

The OPGD board is a new GE P2P optical access board developed for V800R008 and is mainly used for FTTH household user access and for DSLAM convergence. The OPGD board also supports FTTM (mobile bearing) and FTTO (enterprise users).

Purpose

P2P optical access boards prior to OPGD include OPFA, and SPUA.ETHB, The following table lists the ports provided and scenarios supported by each board. Compared with other P2P optical access boards, the OPGD board features more advantages for the access and the subtending scenarios.

Board	Port	Application Scenario
OPFA	16 FE optical ports	It can be directly connected to home user terminal (ONT) only and does not support subtending or upstream transmission. It is connected to the ONT to implement FTTH and provides a 100 Mbit/s bandwidth to each household.
OPGD	48 GE optical ports	It supports the access and subtending scenarios and does not support upstream transmission. • In the access scenario, it is connected to the ONT to implement FTTH and provides a 1000 Mbit/s bandwidth to each household. • In the subtending scenario, it is connected to the DSLAM, CBU, or SBU to implement FTTC/FTTB, FTTO, or FTTM respectively.
ETHB	8 GE optical/electrical ports	It supports subtending and upstream transmission, but cannot be directly connected to home user terminal. • In the subtending scenario, it is connected to the DSLAM to implement FTTC/FTTB. Through the convergence by the DSLAM, each GE port can provide services for a large number of users. • In the upstream transmission scenario, the ETHB board functions as a GIU upstream

Board	Port	Application Scenario
		interface board. It extends the number of upstream ports in the system to increase the total upstream bandwidth of the system.
SPUA	8 GE optical ports+2 10GE optical ports	It supports subtending and upstream transmission, but cannot be directly connected to home user terminal. • In the subtending scenario, it is connected to the DSLAM to implement FTTC/FTTB. Through the convergence by the DSLAM, each GE port can provide services for a large number of users. • In the upstream transmission scenario, it provides a high upstream forwarding bandwidth. It implements upstream link backup by inter-board aggregation and inter-board protect group.

The OPGD board provides GE P2P Ethernet optical access for more flexible FTTx solutions at higher bandwidth, lower costs, and higher reliability.

- Higher bandwidth. Traditional FE P2P optical access provides only a 100 Mbit/s transmission rate, but GE P2P optical access allows for 1000 Mbit/s. The FTTH solution implemented through GE P2P optical access can provide a higher bandwidth for users, thus meeting the requirements of high-end users.
- Lower costs. Compared with SPUA and ETHB, which are capable of both upstream transmission and subtending, the OPGD board is specially designed for subtending and access scenarios. The OPGD board provides 48 GE ports, so it can be subtended to more DSLAMs and hence reduces the costs of FTTC/FTTB networking.
- Higher reliability. The OPGD board allows a higher reliability in the DSLAM subtending scenario through features such as inter-board aggregation, smart link, and ring check.
- More flexible scenarios. The OPGD board coordinates with a variety of downstream devices (such as the DSLAM, ONT, SBU, and CBU) to implement FTTC/FTTB, FTTH, FTTO, and FTTM. An MA5600T/MA5603T/MA5608T configured with the OPGD board can not only be directly connected to access terminals but also subtend DSLAMs in order to converge a large number of users.

Benefit

Benefits to carriers

One MA5600T/MA5603T/MA5608T can support multi-access such as GPON, xDSL, and P2P. Such an All-in-one solution reduces the equipment CAPEX as well as OPEX for carriers.

Benefits to users

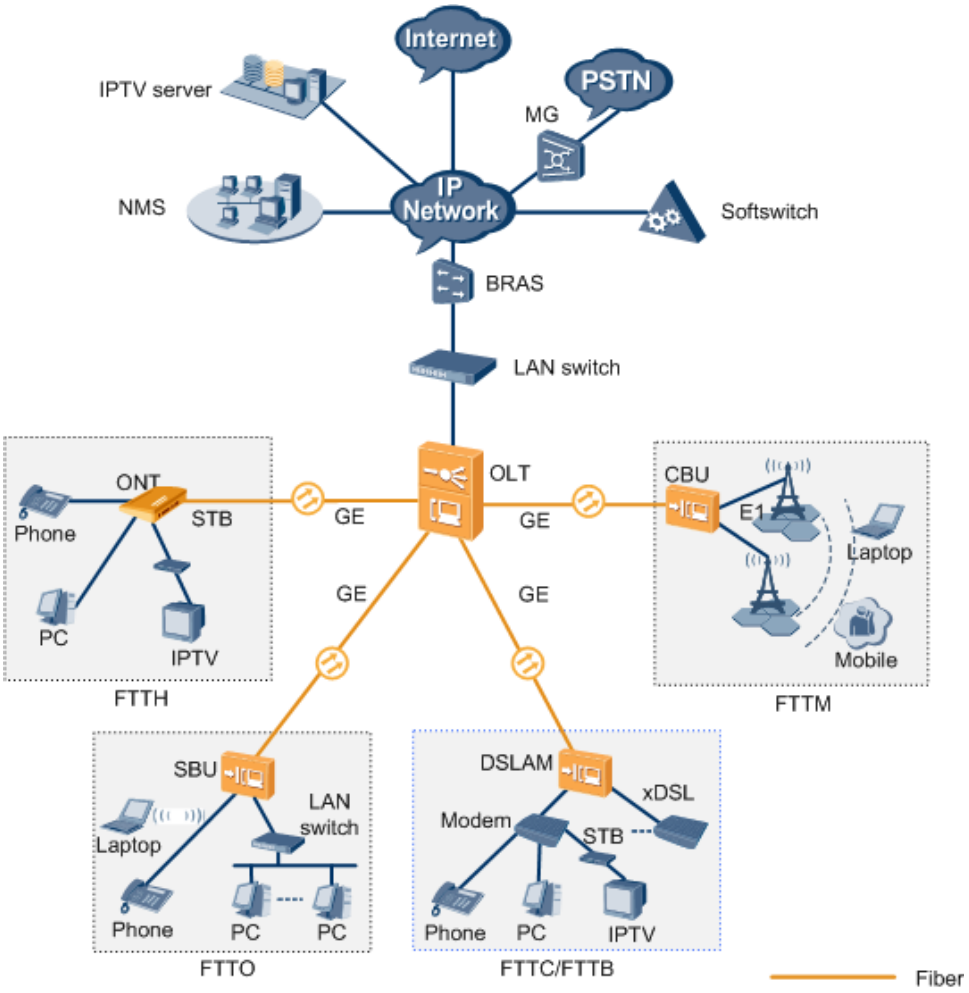
Because the OPGD board can provide high-density GE ports for subtending DSLAMs, which converge massive users, lower costs are needed for providing end-to-end service guarantee for VIP household and enterprise users. In residential communities where optical fibers are

already deployed, a 1000 Mbit/s bandwidth can be provided for high-end users exclusively, meeting the user needs for HD video, voice, and data integrated services.

4.2.2 Network Applications

Figure 4-2 shows the FTTx network application in the GE P2P Ethernet optical access mode.

Figure 4-2 Network application in the GE P2P Ethernet optical access mode



To meet the requirements of different scenarios, the OLT works with ONUs of various types to implement network applications in multiple optical access modes, such as FTTC/FTTB, FTTH, FTTO, and FTTM.

The FTTx network applications in GE P2P Ethernet optical access have the following in common: The data, voice, and video signals of terminal users are sent to ONUs, where the signals are converted into Ethernet packets and then transmitted over optical fibers to the OLT through the GE upstream ports of the ONUs. Then, the Ethernet packets are forwarded to the upper-layer IP network through the upstream port of the OLT.

The differences of the FTTx network applications in GE P2P Ethernet optical access are as follows:

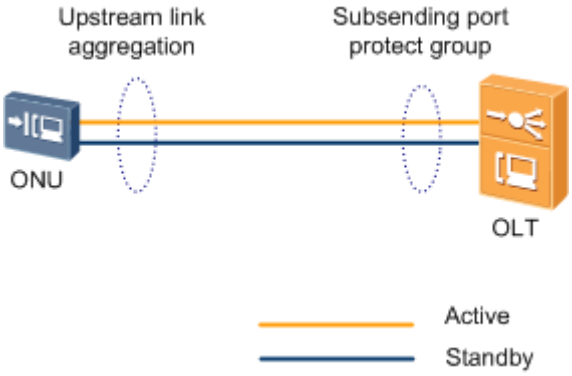
- FTTH: The OLT is connected to the ONUs at user premises through GE P2P Ethernet optical access. In this way, gigabit bandwidth is exclusively provided to each household. FTTH is applicable to new apartments or villas in loose distribution. In this scenario, FTTH provides services of higher bandwidth for high-end users.
- FTTB/FTTC: The OLT is connected to DSLAMs in corridors (FTTB) or by the curb (FTTC) through GE P2P Ethernet optical access. The DSLAMs are then connected to user terminals through xDSL. With the aggregation provided by the DSLAMs, one port on the OPGD board can be connected to a large number of users. FTTB/FTTC is applicable to densely-populated residential communities or office buildings. In this scenario, FTTB/FTTC provides services of certain bandwidth for common users.
- FTTO: The OLT is connected to enterprise SBUs through GE P2P Ethernet optical access. The SBUs are connected to user terminals through FE, POTS, or Wi-Fi. QinQ VLAN encapsulation is implemented on the SBUs and the OLT. In this way, transparent and secure data channels can be set up between the enterprise private networks located at different places, and thus the service data and BPDUs between the enterprise private networks can be transparently transmitted over the public network. FTTO is applicable to enterprise networks. In this scenario, FTTO implements TDM PBX, IP PBX, and private line service in the enterprise intranets.
- FTTM: The OLT is connected to CBUs through GE P2P Ethernet optical access. The CBUs are then connected to wireless base stations through E1. The OLT connects wireless base stations to the core IP bearer network through optical access technologies. This implementation mode is not only simpler than traditional SDH/ATM private line technologies, but also drives down the costs of base station backhaul. FTTM is applicable to reconstruring and capacity expansion of mobile bearer networks. In this scenario, FTTM converges the fixed network and the mobile network on the bearer plane.

Network Protection

FTTC/FTTB, FTTO, and FTTM, compared with FTTH, involve a larger number of access users. Hence, network reliability must be ensured. The ONU provides dual upstream ports to implement link redundancy backup. With the coordination of the ONU, the OPGD board on the OLT supports the following link backup modes: inter-board aggregation, smart link, and monitor link.

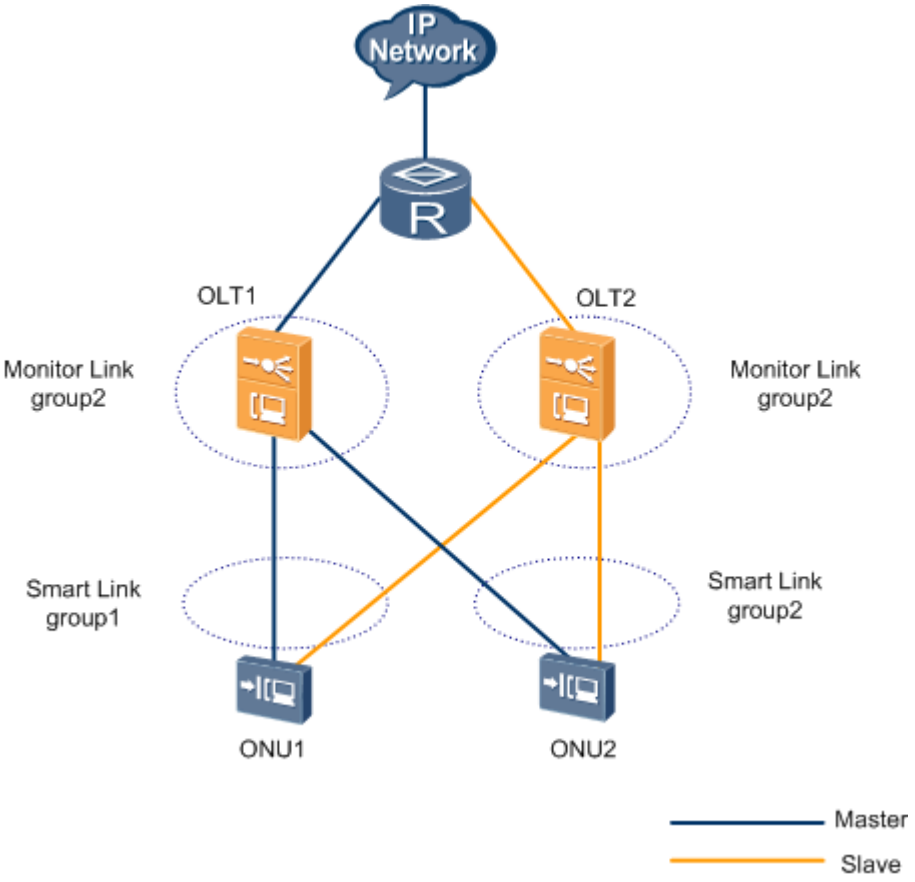
Inter-board aggregation: Two upstream ports of the ONU are respectively connected to two adjacent OPGD boards of the OLT. Dual upstream link aggregation is configured on the ONU, and a protect group is configured on the OLT. Thus, 1:1 backup of GE links can be implemented through inter-board aggregation. Figure 4-3 shows the network topology of the OLT subtending the ONU to implement inter-board aggregation. For more details on the network application of inter-board aggregation, see 19.3.4 Ethernet Link Aggregation Network Applications.

Figure 4-3 Network topology of inter-board aggregation



Smart link and monitor link: Two upstream ports of the ONU are respectively connected to the OPGD board on two OLTs. Monitor link is configured on the OLTs, and smart link is configured on the ONU. 1:1 GE link backup is implemented through a mode similar to type B dual homing of GPON ports. Figure 4-4 shows the network topology of the OLTs subtending the ONUs to implement smart link and monitor link. For more details on smart link and monitor link, see 19.5 Smart Link and Monitor Link.

Figure 4-4 Network topology of smart link and monitor link



4.2.3 Reference Standards and Protocols

The following lists the reference standards and protocols of the OPGD board:

- IEEE 802.3z: 1000Base-SX and 1000Base-LX GE standard
- IEEE 802.1p: Layer 2 service priority QoS and CoS standard
- IEEE 802.1d: standard of MAC bridges
- IEEE 802.1q: VLAN definition standard
- IEEE 802.3x: standard of flow control in full duplex

4.3 Configuring the P2P Optical Fiber Access Service

Point-to-point (P2P) optical access means the point-to-point FTTx access based on the combination between its P2P optical access board and the ONUs. So as to satisfy the users' requirements for the next generation access equipment which integrates video, voice, and data services.

4.3.1 Configuring the FTTH P2P Optical Fiber Access Service (Single-Port for Multiple Services)

Users connected to the OLT through an ONT, and are therefore provided with the Internet, VoIP, and IPTV service through a same port.

Service Requirements

- ONT_1 and ONT_2 are provided with the triple play service through FTTH.
- The Internet access service is provided in the PPPoE access mode.
- The IPTV user connected to ONT_1 can watch all the programs, and the IPTV user connected to ONT_2 can watch only program BTV-1.
- The VoIP service and the IPTV service are provided in the DHCP mode and obtain IP addresses from the DHCP server in the DHCP option-60 mode.
- After receiving different traffic streams, the OLT provides different QoS guarantees to the traffic streams according to the priorities of the traffic streams.
- Traffic streams are differentiated on the OLT by the user-side VLAN (C-VLAN).

Figure 4-5 Example network of the optical fiber access service in the single-port for multiple services mode

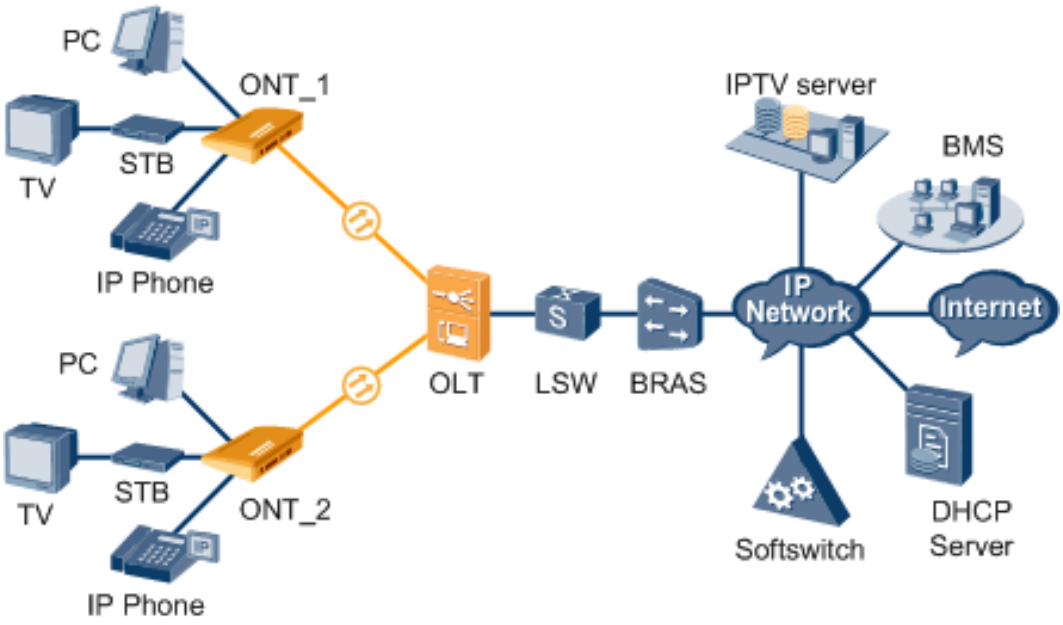


Table 4-1 Data plan for configuring the VLANs

Configuration Item	Data Item	Data
SVLAN	HSI service	SVLAN: 100 CVLAN: 2
	IPTV service	SVLAN: 1000 CVLAN: 4
	VoIP service	SVLAN: 200 CVLAN: 3
IPTV service data	Multicast protocol	IGMP proxy
	Multicast version	IGMP V3
	Configuration mode of the multicast program	Static configuration mode
	IP address of the multicast server	10.10.10.10
	Multicast DHCP server group	20.2.2.2 20.2.2.3
	Multicast program	BTV-1: 224.1.1.10 BTV-2: 224.1.1.20
QoS (priority)	HSI service	Priority: 1; queue scheduling: WRR

Configuration Item	Data Item	Data
	IPTV service	Priority: 4; queue scheduling: WRR
	VoIP service	Priority: 5; queue scheduling: PQ
VoIP service data	VoIP DHCP server group	20.1.1.2 20.1.1.3

Prerequisite

- The OLT is connected to the upper-layer devices such as the BRAS, multicast server, SoftX3000, and DHCP server.
- The VLAN of the LAN switch port connected to the OLT is the same as the upstream VLAN of the OLT.
- The OLT uses the OPFA board or the OPGD board to connect to the ONT.

Procedure

- Configure the Internet access service on the OLT.
  a. Create a VLAN and add an upstream port to the VLAN.

     The VLAN ID is 100, and the VLAN is a smart VLAN. The upstream port is 0/19/0.

     ```
     huawei(config)#vlan 100 smart
     huawei(config)#port vlan 100 0/19 0
     ```

  b. Configure a traffic profile.

     Because the VoIP, IPTV, and Internet access services are provided through the same port, you must set the 802.1p priority of each service. Generally, the priorities are in a descending order for the VoIP service, IPTV service, and Internet access service. In this example, set the traffic profile index to 7 and the priority of the Internet access service to 1.

     ```
     huawei(config)#traffic table ip index 7 cir 10240 priority 1 priority-policy
      local-Setting
     ```

  c. Configure a service port.

     Add a service port to the VLAN and use traffic profile 7. The user-side VLAN ID is 2.

     ```
     huawei(config)#service-port vlan 100 eth 0/5/2 multi-service user-vlan 2
      rx-cttr 7 tx-cttr 7
     huawei(config)#service-port vlan 100 eth 0/5/3 multi-service user-vlan 2
     rx-cttr
      7 tx-cttr 7
     ```

  d. Configure queue scheduling.

     Use the 3PQ+5WRR queue scheduling. Queues 0-4 adopt the WRR mode, with the weights of 10, 10, 20, 20, and 40 respectively; queues 5-7 adopt the PQ mode.

NOTE

Queue scheduling is a global configuration. You need to configure queue scheduling only once on the OLT, and then the configuration takes effect globally. In the subsequent phases, you need not configure queue scheduling repeatedly when configuring other services.

```
huawei(config)#queue-scheduler wrr 10 10 20 20 40 0 0 0
```

Configure the mapping between queues and 802.1p priorities. Priorities 0-7 map queues 0-7 respectively.

```
huawei(config)#cos-queue-map cos0 0 cos1 1 cos2 2 cos3 3 cos4 4 cos5 5 cos6
6
 cos7 7
```

NOTE

For the service board that supports only four queues, the mapping between 802.1p priorities and queue IDs is as follows: priorities 0 and 1 map queue 1; priorities 2 and 3 map queue 2; priorities 4 and 5 map queue 3; priorities 6 and 7 map queue 4.

e. Save the data.

```
huawei(config)#save
```

- Configure the VoIP service on the OLT.

  a. Create a VLAN and add an upstream port to the VLAN.

  The VLAN ID is 200, and the VLAN is a smart VLAN. The upstream port is0/19/0.

```
huawei(config)#vlan 200 smart
huawei(config)#port vlan 200 0/19 0
```

  b. Configure a traffic profile.

  The traffic profile index is 8, and the 802.1p priority of the VoIP service is 6.

```
huawei(config)#traffic table ip index 8 cir 10240 priority 6 priority-policy
 local-Setting
```

  c. Configure a service port.

  Add a service port to the VLAN and use traffic profile 8. The user-side VLAN ID is 3.

```
huawei(config)#service-port vlan 200 eth 0/5/2 multi-service user-vlan 3
 rx-cttr 8 tx-cttr 8
huawei(config)#service-port vlan 200 eth 0/5/3 multi-service user-vlan 3
 rx-cttr 8 tx-cttr 8
```

  d. Configure the DHCP relay.

  The VoIP service and the IPTV service are provided in the DHCP mode. The DHCP option 60 domain is used to differentiate service types.

  - The DHCP domain of the VoIP service is **voice**.
  - The IP addresses of VoIP DHCP server group 1 are 20.1.1.2 and 20.1.1.3.
  - The IP address of the Layer 3 interface of VLAN 200 is 10.1.1.1/24.
  - The gateway IP address of the DHCP domain is 10.1.1.1/24.

```
huawei(config)#dhcp mode layer-3 option-60
huawei(config)#dhcp-server 1 ip 20.1.1.2 20.1.1.3
huawei(config)#dhcp domain voice
huawei(config-dhcp-domain-voice)#dhcp-server 1
huawei(config-dhcp-domain-voice)#quit
huawei(config)#interface vlanif 200
huawei(config-if-vlanif200)#ip address 10.1.1.1 24
```

```
huawei(config-if-vlanif200)#dhcp domain voice gateway 10.1.1.1
huawei(config-if-vlanif200)#quit
```

NOTE

The DHCP option 60 domain of the Ethernet phone (Ephone) varies with the terminal type. In the actual configuration, see the operation instructions of the Ephone.

e. Save the data.

```
huawei(config)#save
```

- Configure the IPTV service on the OLT.

a. Create a VLAN and add an upstream port to the VLAN.

The VLAN ID is 1000, and the VLAN is a smart VLAN. The upstream port is0/19/0.

```
huawei(config)#vlan 1000 smart
huawei(config)#port vlan 1000 0/19 0
```

b. Configure a traffic profile.

The traffic profile index is 9, and the 802.1p priority of the IPTV service is 5.

```
huawei(config)#traffic table ip index 9 cir off priority 5 priority-policy
 local-Setting
```

c. Configure a service port.

Add a service port to the VLAN and use traffic profile 9. The user-side VLAN ID is 4.

```
huawei(config)#service-port 200 vlan 1000 eth 0/5/2 multi-service user-vlan
4
 rx-cttr 9 tx-cttr 9
huawei(config)#service-port 300 vlan 1000 eth 0/5/3 multi-service user-vlan
4
 rx-cttr 9 tx-cttr 9
```

d. Configure the DHCP relay.

The VoIP service and the IPTV service are provided in the DHCP mode. The DHCP option 60 domain is used to differentiate service types.

- The DHCP domain of the IPTV service is **video**.
- The IP addresses of IPTV DHCP server group 2 are 20.2.2.2 and 20.2.2.3.
- The IP address of the Layer 3 interface of VLAN 1000 is 10.2.2.1/24.
- The gateway IP address of the DHCP domain is 10.2.2.1/24.

```
huawei(config)#dhcp mode layer-3 option-60
huawei(config)#dhcp-server 2 ip 20.2.2.2 20.2.2.3
huawei(config)#dhcp domain video
huawei(config-dhcp-domain-video)#dhcp-server 2
huawei(config-dhcp-domain-voice)#quit
huawei(config)#interface vlanif 1000
huawei(config-if-vlanif1000)#ip address 10.2.2.1 24
huawei(config-if-vlanif1000)#dhcp domain video gateway 10.2.2.1
huawei(config-if-vlanif1000)#quit
```

NOTE

The DHCP option 60 domain of the set-top box (STB) varies with the terminal type. In the actual configuration, see the operation instructions of the STB.

e. Create a multicast VLAN and select the IGMP mode.

Select the IGMP proxy mode.

```
huawei(config)#multicast-vlan 1000
huawei(config-mvlan1000)#igmp mode proxy
  Are you sure to change IGMP mode?(y/n)[n]:y
```

f. Set the IGMP version.

Set the IGMP version of the multicast VLAN to IGMP v3.

```
huawei(config-mvlan1000)#igmp version v3
```

g. Add an IGMP upstream port.

The IGMP upstream port is port 0/19/0 and works in the default mode, and protocol packets are transmitted to all the IGMP upstream ports in the multicast VLAN.

```
huawei(config-mvlan1000)#igmp uplink-port 0/19/0
huawei(config-mvlan1000)#btv
huawei(config-btv)#igmp uplink-port-mode default
Are you sure to change the uplink port mode?(y/n)[n]:y
```

h. (Optional) Set the multicast global parameters.

In this example, the default settings are used for all the multicast global parameters.

i. Configure the program library.

Configure the program names to BTV-1 and BTV-2, multicast IP addresses of the programs to 224.1.1.10 and 224.1.1.20, and source IP address of the programs to 10.10.10.10.

```
huawei(config-btv)#multicast-vlan 1000
huawei(config-mvlan1000)#igmp program add name BTV-1 ip 224.1.1.10 sourceip
 10.10.10.10
huawei(config-mvlan1000)#igmp program add name BTV-2 ip 224.1.1.20 sourceip
 10.10.10.10
```

j. Configure the right profile.

Configure the profile name to profile0, with the right of watching program BTV-1.

```
huawei(config-mvlan1000)#btv
huawei(config-btv)#igmp profile add profile-name profile0
huawei(config-btv)#igmp profile profile-name profile0 program-name BTV-1 watch
```

k. Configure the multicast users.

Add service ports 200 and 300 as multicast users.

```
huawei(config-btv)#igmp user add service-port 200 no-auth
huawei(config-btv)#igmp user add service-port 300 auth
huawei(config-btv)#igmp user bind-profile service-port 300 profile-name
profile0
huawei(config-btv)#multicast-vlan 1000
huawei(config-mvlan1000)#igmp multicast-vlan member service-port 200
huawei(config-mvlan1000)#igmp multicast-vlan member service-port 300
huawei(config-mvlan1000)#quit
```

l. Save the data.

```
huawei(config)#save
```

----End

Result

After the related upstream device and downstream device are configured, the triple play service (Internet, VoIP, and IPTV services) is available.

- The Internet user can access the Internet in the PPPoE mode.
- The VoIP user can make and receive phone calls.
- The IPTV user connected to port 0/5/2 can watch all the programs, and the IPTV user connected to port 0/5/3 can watch only program BTV-1.

Configuration File

Internet service:

```
vlan 100 smart
port vlan 100 0/19 0
traffic table ip index 7 cir 10240 priority 1 priority-policy local-Setting
service-port vlan 100 eth 0/5/2 multi-service user-vlan 2 rx-cttr 7 tx-cttr 7
service-port vlan 100 eth 0/5/3 multi-service user-vlan 2 rx-cttr 7 tx-cttr 7
queue-scheduler wrr 10 10 20 20 40 0 0 0
cos-queue-map cos0 0 cos1 1 cos2 2 cos3 3 cos4 4 cos5 5 cos6 6 cos7 7
save
```

VoIP service:

```
vlan 200 smart
port vlan 200 0/19 0
traffic table ip index 8 cir 10240 priority 6 priority-policy local-Setting
service-port vlan 200 eth 0/5/2 multi-service user-vlan 3 rx-cttr 8 tx-cttr 8
service-port vlan 200 eth 0/5/3 multi-service user-vlan 3 rx-cttr 8 tx-cttr 8
dhcp mode layer-3 option-60
dhcp-server 1 ip 20.1.1.2 20.1.1.3
dhcp domain voice
dhcp-server 1
quit
interface vlanif 200
ip address 10.1.1.1 24
dhcp domain voice gateway 10.1.1.1
quit
save
```

IPTV service:

```
vlan 1000 smart
port vlan 1000 0/19 0
traffic table ip index 9 cir off priority 5 priority-policy local-Setting
service-port 200 vlan 1000 eth 0/5/2 multi-service user-vlan 4 rx-cttr 9 tx-cttr 9
service-port 300 vlan 1000 eth 0/5/3 multi-service user-vlan 4 rx-cttr 9 tx-cttr 9
dhcp mode layer-3 option-60
dhcp-server 2 ip 20.2.2.2 20.2.2.3
dhcp domain video
dhcp-server 2
quit
interface vlanif 1000
ip address 10.2.2.1 24
dhcp domain video gateway 10.2.2.1
quit
```

```
multicast-vlan 1000
igmp mode proxy
y
igmp uplink-port
igmp program add name BTV-1 ip 224.1.1.10 sourceip 10.10.10.10
igmp program add name BTV-2 ip 224.1.1.20 sourceip 10.10.10.10
btv
igmp uplink-port-mode default
y
igmp profile add profile-name profile0
igmp profile profile-name profile0 program-name BTV-1 watch
igmp user add service-port 200 no-auth
igmp user add service-port 300 auth
igmp user bind-profile service-port 300 profile-name profile0
multicast-vlan 1000
igmp multicast-vlan member service-port 200
igmp multicast-vlan member service-port 300
quit
save
```

4.3.2 Configuring MDUs Subtended to an OLT

MDUs are subtended to an OLT through the OPGD board, thereby saving upstream optical fibers and simplifying the network and service configuration.

Service Requirements

- MDU_1 and MDU_2 are connected to an OLT through GE subtending, implementing the Internet access service.
- The Internet access service is provided in the PPPoE dialing mode.

Figure 4-6 Network of MDUs subtended to an OLT

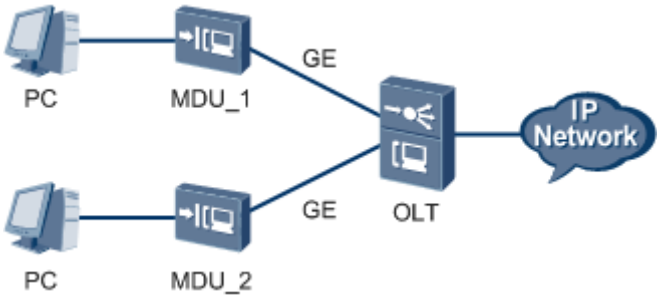


Table 4-2 Data plan

Item	Data
OLT	SVLAN ID: 100 SVLAN type: smart VLAN
	CVLAN ID: 200
	Upstream port: 0/19/0

Item	Data
MDU_1	SVLAN ID: 200 SVLAN type: smart VLAN
	Upstream port: 0/0/1 **NOTE** The upstream ports vary with MDU type.
MDU_2	SVLAN ID: 200 SVLAN type: smart VLAN
	Upstream port: 0/0/1

Procedure

- Configure the OLT.
  a. Configure the port role.

     Configure the port role of the OPGD board as a subtending port. The port roles of the OPGD board are user port and subtending port. By default, the port role is user port.

     ```
     huawei(config)#interface opg 0/2
     huawei(config-if-opg-0/2)#network-role cascade
     huawei(config-if-opg-0/2)#quit
     ```

  b. Create a VLAN and add an upstream port to the VLAN.

     Create smart SVLAN 100. The upstream port is port 0/19/0.

     ```
     huawei(config)#vlan 100 smart
     huawei(config)#port vlan 100 0/19 0
     ```

  c. Configure a service port.

     Add the service port to the SVLAN by using default traffic profile 6. The CVLAN ID is 200, the same as the upstream VLAN ID of the MDU. MDU_1 and MDU_2 are connected to ports 0/2/0 and 0/2/1 of the OLT respectively.

     ```
     huawei(config)#service-port vlan 100 eth 0/2/0 multi-service user-vlan 200
      rx-cttr 6 tx-cttr 6
     huawei(config)#service-port vlan 100 eth 0/2/1 multi-service user-vlan 200
      rx-cttr 6 tx-cttr 6
     ```

  d. Save the data.

     ```
     huawei(config)#save
     ```

- Configure the MDUs.

  The configurations of MDU_1 and MDU_2 are the same. The configuration of MDU_1 is used as an example.

  a. Create a VLAN and add an upstream port to the VLAN.

     Create smart SVLAN 200. The upstream port is port 0/0/1.

  NOTE
  The SVLAN of the MDU must be the same as the CVLAN of the OLT.

  ```
  huawei(config)#vlan 200 smart
  huawei(config)#port vlan 200 0/0 1
  ```

b. Configure a service port.

According to actual conditions, an MDU supports multiple access modes. In this example, the ethernet port 0/3/1 is used. For other access modes, see the corresponding configuration guide of the MDU.

```
huawei(config)#service-port vlan 200 eth 0/3/1 multi-service user-vlan untagged
 rx-cttr 6 tx-cttr 6
```

c. Save the data.

```
huawei(config)#save
```

----End

Result

On the PC, the Internet access service is provided in the PPPoE dialing mode.

Configuration File

Configure the OLT:

```
interface opg 0/2
network-role cascade
quit
vlan 100 smart
port vlan 100 0/19 0
service-port vlan 100 eth 0/2/0 multi-service user-vlan 200 rx-cttr 6 tx-cttr 6
service-port vlan 100 eth 0/2/1 multi-service user-vlan 200 rx-cttr 6 tx-cttr 6
save
```

Configure the MDU:

```
vlan 200 smart
port vlan 200 0/0 1
service-port vlan 200 eth 0/3/1 multi-service user-vlan untagged rx-cttr 6 tx-cttr 6
save
```

5 ADSL2+ Access

About This Chapter

ADSL2+ is an extension of ADSL. The maximum upstream and downstream rates reach 2.5 Mbit/s and 24 Mbit/s, respectively. The maximum transmission distance is 6.5 km.

5.1 ADSL2+ Access Introduction

Definition

ADSL2+ is an extension of ADSL, an asymmetric transmission technology that transmits data at a high speed over twisted pairs.

Table 5-1 lists the comparisons between technical specifications of ADSL, ADSL2, ADSL2+, and VDSL2.

Table 5-1 Comparisons between technical specifications of ADSL, ADSL2, ADSL2+, and VDSL2

Technology	Operating Frequency (Hz)	Upstream and Downstream Rate (bit/s)	Maximum Transmission Distance (km)
ADSL	26 k to 138 k 138 k to 1.1 M	896 k/8196 k	5
ADSL2	26 k to 138 k 138 k to 1.1 M	1.2 M/12 M	5.2
ADSL2+	26 k to 138 k 138 k to 2.2 M	2.5 M/24 M	6.5
VDSL2	Stop frequency: 30 M	40 M/80 M	3.5

Purpose

The ADSL2+ technology supports asymmetric transmission in upstream and downstream directions, which provides high-speed data transmission for services. The ADSL2+ supports:

- Backward compatibility. ADSL2+ is compatible with ADSL and ADSL2. Compared with ADSL, both ADSL2 and ADSL2+ are improved from the aspect of not only transmission distance, line outgoing ratio (Number of line pairs on which ADSL services can be provisioned/Total number of line pairs in a feeder cable bundle), and downstream bandwidth but also power management and fault detection. The new functions and features supported by ADSL2 and ADSL2+ improve network performance and cooperation capability. Accordingly, carriers can deploy ADSL2 or ADSL2+ services by upgrading existing devices, which implements new applications and services with reduced costs.
- Forward evolution. ADSL2+ is compatible with VDSL2 and provides a longer transmission distance than VDSL2.

ADSL2 or ADSL2+ applies in the area which requires a long transmission distance (longer than 1.2 km) and a high bandwidth. ADSL2+ applies in the area with low line outgoing ratio. This reduces interference between line pairs and improves the line outgoing ratio.

5.2 Basic ADSL2+ Technologies

5.2.1 Spectrum Plan

The factors affecting DSL loops may vary depending on network conditions, and it is difficult to address the application requirements of different scenarios using a single mechanism. To account for this, plan spectra to form various spectrum profiles for various usage scenarios. Select a proper Annex type and PSD profile to configure a spectrum profile.

NOTE

Knowledge about the G.993.2, G.997.1, and TR-165 standards helps you better understand the spectrum plan described in this section.

NOTE

ADSL2+ line parameters can be used in different combinations based on profiles. The configuration modes can be classified as RFC2662 (also called the common mode), RFC4706 (also called NGADSL mode), and TR165. For the MA5600T/MA5603T/MA5608T, the default configuration mode is RFC2662. Carriers can switch between the configuration modes by running the **switch adsl mode to** command. Considering the current development trend, it is recommended that you use TR165, which is more flexible than the others. Unless otherwise specified, the command parameters included in the following ADSL2+-related topics are specific to the TR165 mode.

5.2.2 Annex Type

Most DSL standards provide a generic definition in the body, and then a description about specific schemes in the Annex. The schemes specify how to use the low frequency band in typical application scenarios. The schemes also specify how to plan the upstream/downstream band (apart from the low frequency band) for data transmission and how to plan the power spectrum.

Users can select a proper Annex type by running commands. When an Annex type is selected, the upstream/downstream band plan and power spectrum plan are determined.

NOTE

The power spectrum plan is critical for controlling the performance and reliability of DSL lines. VDSL2 provides flexible power spectrum control mechanisms. The concepts and features related to the power spectrum plan are described in 6.3.6 PSD Profiles. As an Annex type includes a power spectrum plan, this section will also include information about power spectrum. It is recommended that you also read 6.3.6 PSD Profiles to better understand the VDSL2 feature.

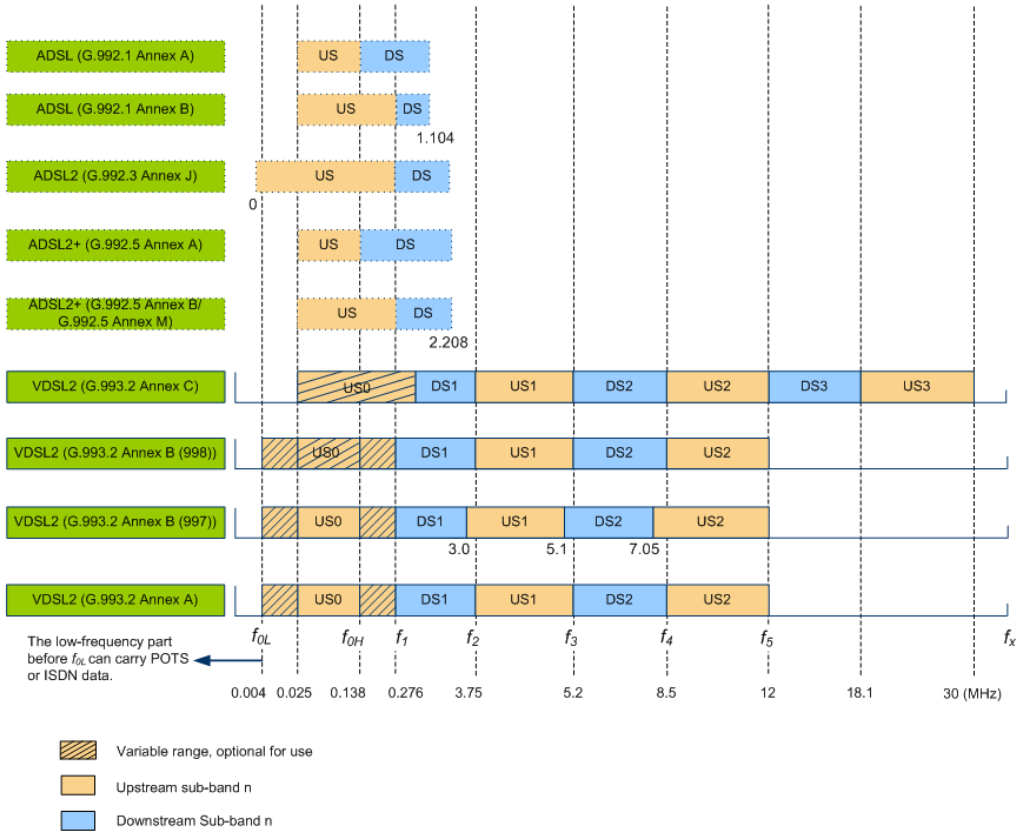
Annex Types and Upstream/Downstream Band Plans

An Annex type defines the scheme for using the low frequency band (the frequency band before f_{0L} as shown in Figure 5-1, used for carrying POTS or ISDN data) and the scheme for planning the upstream/downstream band (apart from the low frequency band) for data transmission. The upstream/downstream band plan specifies the spectral segments for upstream/downstream transmission, and the start and stop frequencies in each segment.

The spectral segment used for upstream transmission is called upstream sub-band (US), such as US0 and US1 in Figure 5-1; the spectral segment used for downstream transmission is called downstream sub-band (DS), such as DS1 and DS2 in Figure 5-1. The total number of USs and DSs in the entire band is the total number of bands specified in the spectrum profile. For example, "5 Band" indicates that the entire band is divided into five sub-bands.

For ADSL/ADSL2/ADSL2+, the entire available spectrum is divided into one US and one DS, as shown in Figure 5-1. This figure also shows mapping between US0 for VDSL2 and US for ADSL2+. The mapping is also described in 6.3.7 Limit PSD Mask.

Figure 5-1 ADSL2+/VDSL2 upstream/downstream band plan



Among the upstream sub-bands, US0 is optional (as shown in Figure 5-1) and an Annex type defines the frequency range of US0 (start frequency f_{0L}; stop frequency f_{0H}) and usage of US0. A Huawei access device also provides commands for enabling and disabling US0 and specifying a PSD mask.

For long-distance access, the upstream high frequency band is fully exploited, so the low frequency band becomes a valuable resource. Enabling US0 in this case will extend the DSL coverage and improve upstream line performance. VDSL2 can be activated beyond 1.4 km only when US0 is enabled. Usually, you are recommended to enable US0 beyond 800 m.

5.2.3 PSD Profiles

The power spectrum plan is a PSD profile that defines the PSD masks for the upstream/downstream frequency bands. In this document, PSD profiles include PSD-related profiles in mode-specific PSD profiles ("mode" refers to transmission mode) and line spectrum profiles defined by TR165.

1. PSD refers to the differential coefficient of the transmit power at the frequency point and reflects the power intensity (expressed in dBm or Hz) at each frequency point. Users can derive the transmit power used in a spectrum band by performing integral calculation for PSD at each frequency point in the band. Controlling the PSD of an ADSL2+ line protects the line against external noise and reduces the interference output of the line.
2. PSD mask is a fold line that links the **maximum PSD** at each frequency point. The system specifies PSD values for a series of breakpoints on a spectrum band and outlines the PSD mask of the spectrum band through an interpolation algorithm.

5.2.4 MIB PSD mask

ITU-T Recommendation G.993.2 defines management information base (MIB)-controlled power spectrum density (PSD) masks for flexible control over PSD.

"MIB-controlled" means configuring PSD masks through the network management system (NMS) or through a digital subscriber line access multiplexer (DSLAM). MIB-controlled PSD masks provide users with more options than the limit PSD masks defined in standards. Carriers can control the power spectrum and reduce crosstalk by configuring suitable PSD masks according to DSLAM distribution, distance to users, and coexistence of ADSL and VDSL. Such user-configured PSD masks are referred to as MIB PSD masks.

For details on MIB PSD masks, see MIB-controlled PSD Mask.

5.3 Key ADSL2+ Techniques

5.3.1 Key Techniques for Improving Line Protection

DSL provides various techniques for improving line protection, such as enhanced error detection and correction, reserved noise margin, and online reconfiguration (OLR). All the techniques employed translate into higher line stability.

Interleaving FEC

Forward error correction (FEC), though having powerful error correction capability, is insufficient for handling long strings of consecutive bit errors that are generated in severe line

noise. Hence, interleaving FEC is introduced. Interleaving FEC is a major approach for avoiding pulse interference.

Working Principle of Interleaving FEC

Interleaving may be block interleaving or convolutional interleaving, and DSL uses the latter. Compared with convolutional interleaving, block interleaving is simple but less effective. The following uses block interleaving as an example to illustrate the interleaving process.

Figure 5-2 shows a typical interleaver. In this example, the rectangle block refers to an interleaving block and the numbers in the block indicate the sequence in which bits enter the interleaver. Generally, bits are written by row and read by column. The interleaving depth (D) is 3 and interleaving width (I) is 7. In practical applications, an interleaver has greater D and I values.

NOTE

ADSL directly uses the FEC codeword N_{FEC} as the interleaver width, whereas VDSL2 uses the fraction ($I = N_{FEC}/q$) of N_{FEC} as the interleaver width, with q ranging from 1 to 8.

Figure 5-2 Working principle of the interleaver

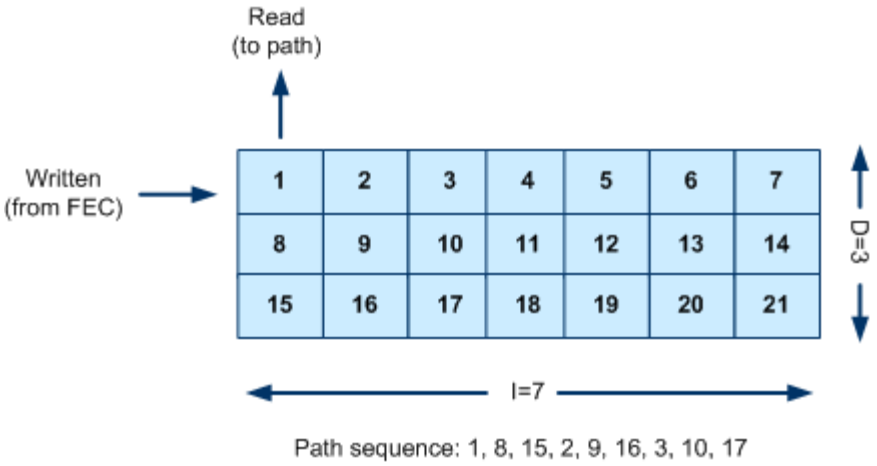


Figure 5-3 shows a de-interleaver that corresponds to the interleaver shown in Figure 5-2. The de-interleaver outputs cells in their correct sequence.

Figure 5-3 Working principle of the de-interleaver

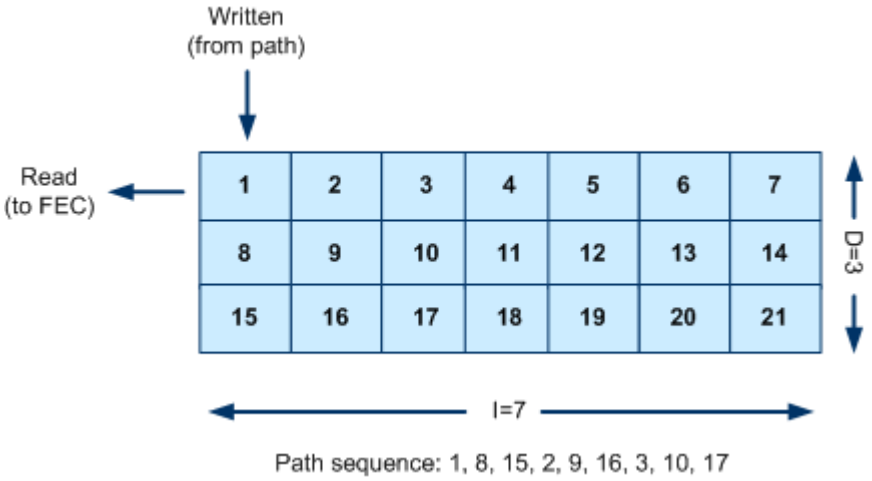


Figure 5-4 shows the benefit of interleaving by comparing the received bit errors with and without interleaving. In the figure, the first two rows indicate the sequence in which bits are transmitted over channels and the last two rows indicate the received bits. If a burst error similar to the third row occurs, bit errors will be distributed when interleaving takes effect so that they can be better corrected.

Figure 5-4 Comparison of received bit errors with and without interleaving

BITS sequence (without interleaving)	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
BITS sequence (with interleaving)	1	8	15	2	9	16	3	10	17	4	11	18	5	12	19	6
the burst errors																
the received BITS (without interleaving)	1	2				6	7	8	9	10	11	12	13	14	15	16
the received BITS (de-interleaving)	1		3	4	5	6	7	8		10	11	12	13	14		16

ITU-T Recommendation G.993.2 also defines a mechanism for dynamically adjusting the interleaving depth (D). In the handshake process, the office and user devices negotiate whether to support dynamic adjustment of the interleaving depth. If yes, the system adjusts the interleaving depth based on line conditions, thereby extending the range for SRA.

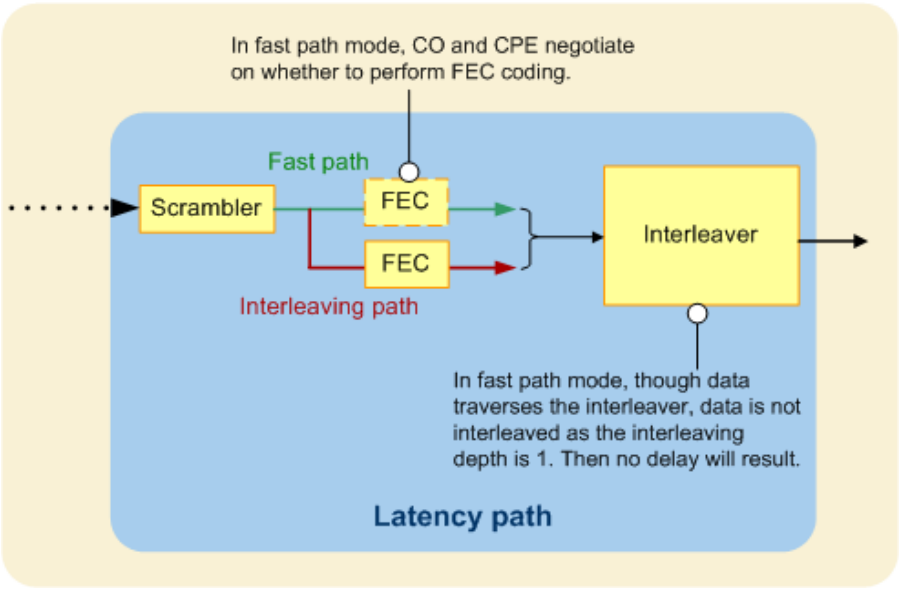
Path Mode and Maximum Interleaving Delay

Interleaving improves the line error correction capability by splitting consecutive bit errors on a line among various FEC frames. As the interleaving takes additional time, delay (referred to as interleaving delay) results. The **maximum interleaving delay** parameter is designed on the MA5600T/MA5603T/MA5608T to control the interleaving delay. Specifically, the interleaving delay produced after a port is activated cannot exceed the maximum interleaving delay. On the MA5600T/MA5603T/MA5608T, users can run the **xdsl inp-delay-profile add** command to set the maximum interleaving delay.

As interleaving delay will impact delay-critical services, such as VoD, voice, and fax services, ADSL2+ allows users to select a path mode ("path" means "latency path" and has the same

meaning as the path in "dual-latency path") before line initialization: fast path or interleaving path. Figure 5-5 shows how the two path modes vary from each other.

Figure 5-5 Fast path and interleaving path



- Fast path: The line has a shorter delay but smaller error correction capability. In this mode, the interleaving depth is 1, which means no interleaving is performed, and the maximum interleaving delay is 0 ms.

NOTE

- ITU-T Recommendation G.997.1 defines three special values for the maximum interleaving delay:
- S0: **Interleaving delay** is set to **0**, indicating no limit on the maximum interleaving delay.
- S1: **Interleaving delay** is set to **1**, indicating the interleaving depth (D) of 1 and the maximum interleaving delay of 0 ms.
- S2: **Interleaving delay** is set to **255**, indicating the maximum interleaving delay of 1 ms.

- Interleaving path: In interleaving path mode, the system has stronger error correction capability but a longer delay. It is typically applicable to the services that are not reliability or delay-critical, such as file download. In this mode, the FEC-processed bit stream is sent to the interleaver and then to the line. On the other side of the line, the bit stream is de-interleaved.

In practical application, the system does not judge the minimum INP or maximum interleaving delay but applies the settings to a board directly. The board will make adaptation to ensure successful line activation after receiving the settings. Generally, use a longer interleaving delay (63 ms, for instance) if the minimum INP value is large (16, for instance). If the minimum INP value is small and the maximum interleaving delay is short, the line will be activated with a low rate or probably cannot be activated.

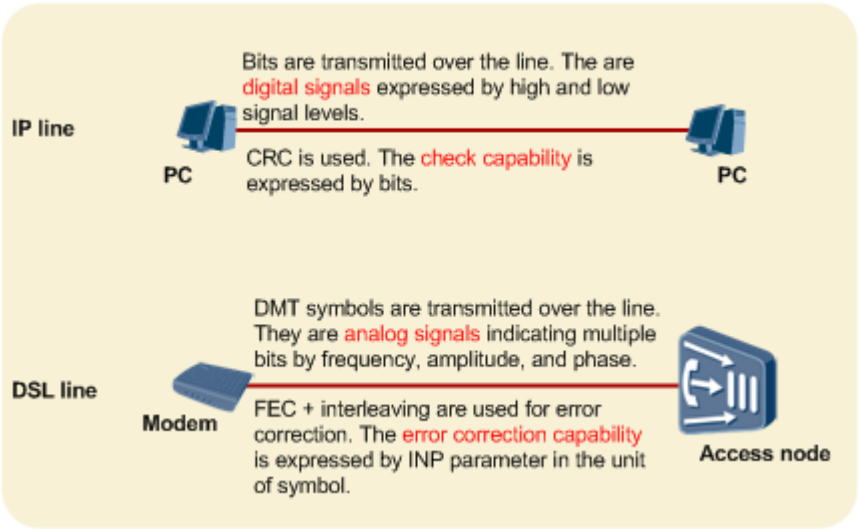
Configurable INP Parameters

Impulse noise protection (INP) refers to a technical category. In the DSL standard, INP indicates the error correction capability of a line or, more specifically, the count of correctable consecutive discrete multi-tone (DMT) symbols during de-interleaving.

INP Definition

Figure 5-6 shows the definition of INP parameters. On the device, **minimum INP** controls the error correction capability. The INP value of an activated port must be greater than or equal to **minimum INP**.

Figure 5-6 INP indication



INP Parameter Application

In ADSL2+/VDSL2, the interleaving capability is represented by **interleaving depth (D)**, the error correction capability by **minimum INP**, and interleaving delay by **maximum interleaving delay** (see Interleaving FEC for details on interleaving), which are correlated to each other. In other words, deeper interleaving means more powerful error correction capability (a greater INP value) but a longer interleaving delay. The three parameters fit a formula defined in ITU-T Recommendation G.993.2.

On the Huawei access device, users can run the **xdsl inp-delay-profile add** command to configure INP (or the interleaving delay). A board adjusts the interleaving depth and delay based on the specified minimum INP for the system to suppress pulse noise interference. If erasure decoding is used, INP can be significantly increased without additional redundancy (no impact on the efficiency for carrying payload).

In practical application, the system does not judge the minimum INP or maximum interleaving delay before applying the settings to a board. The board will make adaptation to ensure successful line activation after receiving the settings. Generally, use a longer interleaving delay (63 ms, for instance) if the minimum INP value is large (16, for instance). If the minimum INP value is small and the maximum interleaving delay is short, the line will be activated with a low rate or probably cannot be activated. This means that there is a correlation between INP and the activated line rate. When the interleaving depth is constant, a greater INP value means a sharper decrease of the activated line rate.

When configuring the minimum INP, users must note the following conditions:

- If the Internet access rate is low, the line probably has a long delay. The most possible cause of the long delay is a large INP value.

- In the ADSL2+/VDSL2 over POTS service, there will be an abrupt change in line impedance after an onhook, producing transient pulse signals on the line. In this case, the ADSL2+/VDSL2 line will lose packets or even result in offline instances. It is recommended to set the minimum INP to 2 or greater for ADSL2+/VDSL2 over POTS.

The optimal INP value must be determined based on statistics of line noise distribution and spectrum range monitored over a long duration in order for the system to minimize the impact on line performance while maintaining a stable line. Impulse noise monitor (INM) is used for the monitoring.

Erasure Decoding

When used with FEC (Reed-Solomon coding), erasure decoding increases the system INP value without requiring additional redundancy.

Erasure decoding is optional as defined in the standard and the device manufacturers decide whether to implement it on central office (CO) and customer premises equipment (CPE) devices.

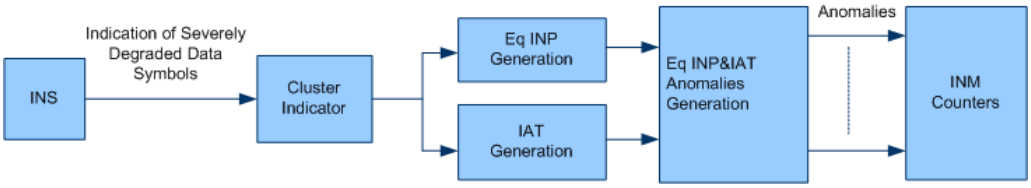
Impulse Noise Monitor (INM)

A greater INP value means more powerful line error correction capability, but longer data transmission delay and lower efficiency of carrying payload. Therefore, setting an optimal INP value is important to ADSL2+/VDSL2.

The optimal INP value must be determined based on statistics of line noise distribution and spectrum range monitored over a long duration in order for the system to minimize the impact on line performance while maintaining a stable line. Impulse noise monitor (INM) is used for the monitoring.

Figure 5-7 shows the working principle of INM.

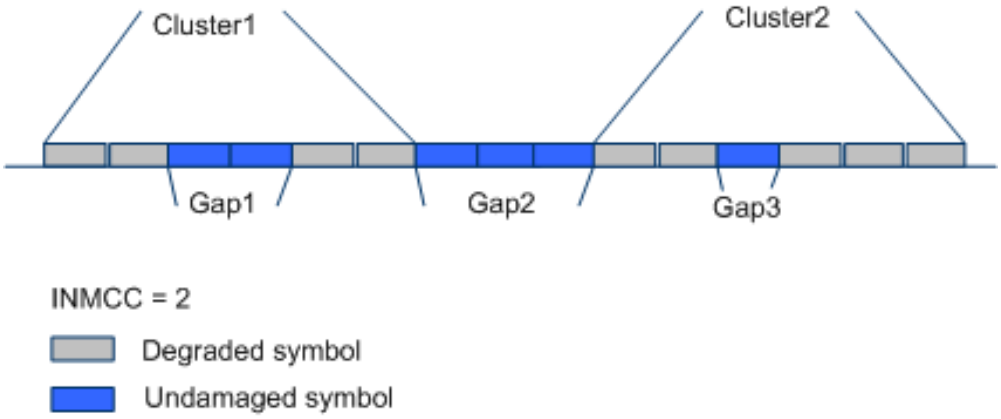
Figure 5-7 Working principle of INM



Working principle of INM:

1. The impulse noise sensor (INS) checks for severe damage in DMT symbols. If DMT symbols are severely damaged, they are downgraded.
2. The cluster indicator identifies INS-detected DMT symbols and groups the matched DMT symbols in a cluster. Clusters are preconditions for later DMT symbol processing. Figure 5-8 shows the process of identifying DMT symbols in clusters.

Figure 5-8 Working principle of INM



- As shown in the figure above, INM cluster continuation value (INMCC) is a key parameter for a cluster. INMCC indicates the maximum number of intact DMT symbols that can be included in a cluster. In this example, INMCC is 2 and Gap1 has two DMT symbols, which belong to a cluster (Cluster 1). Gap2 has three DMT symbols, higher than the limit. Therefore, Cluster1 includes only Gap1 and Gap2 does not belong to any cluster.

3. The Eq INP generation module calculates equivalent INP (INP_eq) for each cluster, and the inter arrive time (IAT) generation module calculates IAT for the entire symbol series. IAT refers to the number of symbols between two consecutive clusters, excluding the Sync symbol.
4. The Eq INP & IAT anomalies generation module collects statistics of **INP_eq** and **IAT**.
5. The INM counters count **INP_eq** and **IAT** by a certain rule, and produce irregular INP_eq and IAT bar charts based on the data. Users can view and use the data, and configure **INP_Min** (minimum INP) and **Delay_Max** (maximum interleaving delay) based on **INP_eq** and **IAT**.
6. Users can query the INM statistical results by running the **display statistics performance** command, or view the INP_eq and IAT bar charts using the NMS.

Physical Layer Retransmission (G.INP)

Some pulse noise may produce numerous bit errors. To protect a system against the pulse noise, one theoretical approach is to improve impulse noise protection (INP) by increasing forward error correction (FEC) redundancy and interleaving depth. However, the theoretical approach is not feasible because it causes a long delay and low efficiency in carrying payload, or has high requirements on components. ITU-T Recommendation G.998.4 defines physical layer retransmission to provide an alternative for improving INP. Specifically, physical layer retransmission improves INP while providing a high transmission rate and an acceptable transmission delay, and it is typically applicable to line quality-critical services, such as video services.

G.INP is another designation of ITU-T Recommendation G.998.4. Physical layer retransmission is referred to as RTX.

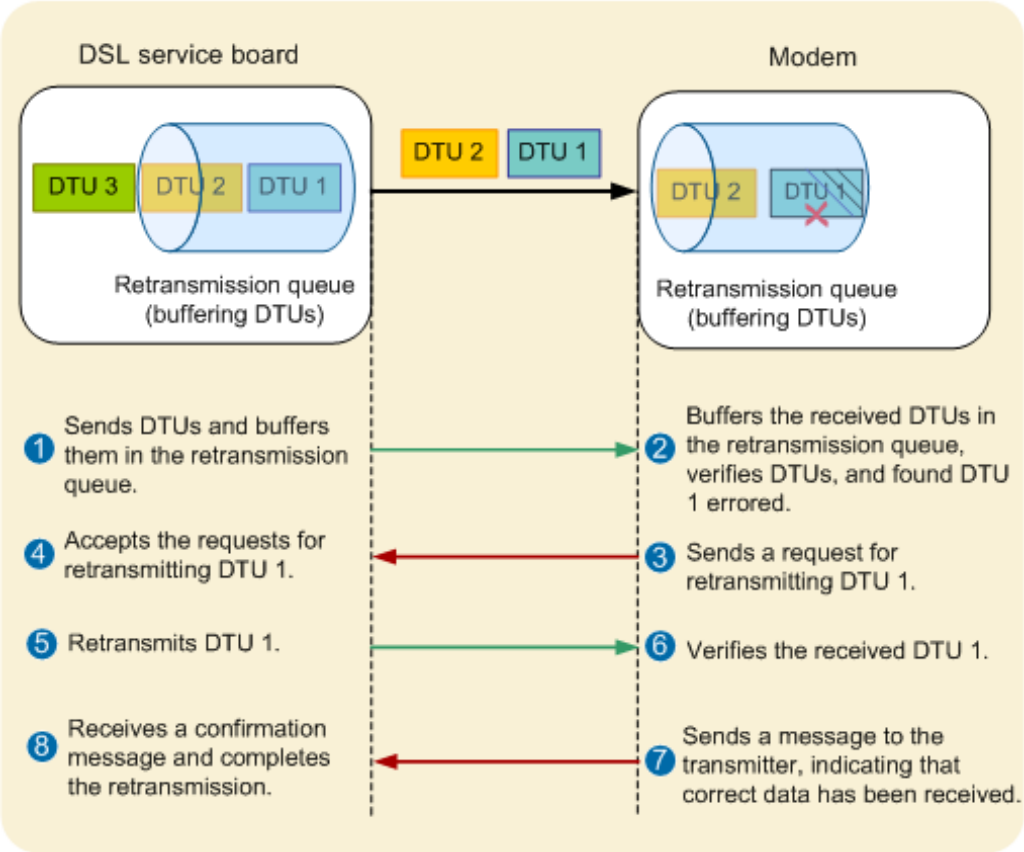
G.INP is intended to protect the system against the following types of pulse noise:

- Single high impulsive noise event (SHINE), which is neither repetitive nor periodic, but unpredictable because it is caused by burst impulse.

- Repetitive electric impulsive noise (REIN), which is repetitive and is caused by the electric main line and influenced by the local AC frequency.

Figure 5-9 shows how the access device implements retransmission in the downstream direction. Retransmission in the upstream direction is similar.

Figure 5-9 Working principle of retransmission



As shown in Figure 5-9, both the transmitter and receiver provide retransmission queues. To start the retransmission process, the transmitter encodes the to-be-sent data in data transfer units (DTUs), which are buffered in a retransmission queue. After receiving the DTUs, the receiver also buffers them in a retransmission queue and verifies them. If a DTU is found errored, the receiver sends a retransmission request to the transmitter. Then, the transmitter retransmits the DTU as requested. When receiving the retransmitted DTU, the receiver verifies it. If the DTU is correct, the receiver sends an acknowledgement message to the transmitter. By now, the retransmission process is completed.

In line with ITU-T Recommendation G.998.4, the Huawei access device supports G.INP retransmission parameter settings. For details, see G.998.4-related parameters in the **xdsl line-spectrum-profile add**, **xdsl inp-delay-profile add**, and **xdsl data-rate-profile add** commands. Users can query statistics of retransmission performance and operation specifications by running the **display xdsl statistics performance**, **display line operation**, and **display channel operation** commands.

Configurable Noise Margin

Noise margin is also signal-to-noise ratio (SNR) margin. The line conditions, such as ambient temperature, humidity, and ambient background noise, keep changing, and so does the SNR of each tone. A noise margin is retained when bits are allocated to each tone. When the line conditions change, the SNR decreases. If the SNR decrease is within the noise margin, the bit error ratio (BER) can stay lower than the standard-stipulated 10^{-7}, and data can be properly transmitted.

Concepts

Noise margin

Noise margin refers to the extra noise that the access device can tolerate while retaining the existing rate and BER. A wider noise margin means a more stable line but a lower activated physical connection rate.

Bit allocation

The noise power spectrum and line attenuation vary with the frequency, and different tones have varied SNRs and number of allocated bits. Therefore, different tones have varied noise margins but only one noise margin value is displayed. In practical application, the lowest noise margin will apply as the noise margin of the entire xDSL line.

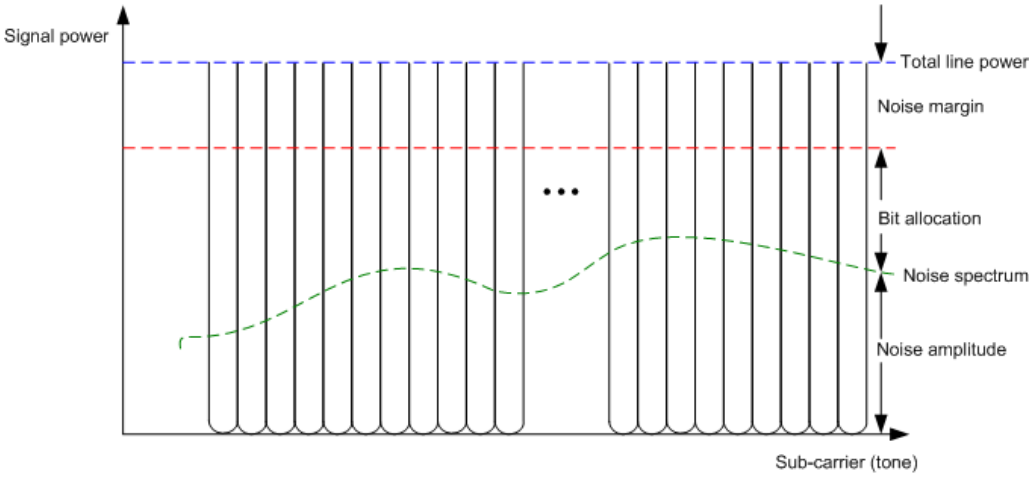
SNR

As a basic indicator in the communication industry, SNR reflects path quality. SNR refers to the ratio of the energy of data signals carried over each tone to the noise energy. Therefore, the xDSL SNR is the SNR of each tone. Each tone's signal and noise energy is expressed in dBm/Hz. Noise power ranges from -120 dBm/Hz to -140 dBm/Hz, and signal transmit power ranges from -40 dBm/Hz to -90 dBm/Hz. A tone with a 3 dB SNR can carry one bit. For a tone to carry 15 bits, the tone must have an SNR of at least 45 dB.

Working Principle

Figure 5-10 shows how noise margin works. Each tube represents a tone, the blue line represents total line power, the area outlined by the blue and red lines represents the reserved noise margin, and the area below the green line represents noise power. As shown in the figure, the area outlined between the red and green lines is used for carrying transmission signals (bit allocation).

Figure 5-10 Noise margin



When no noise margin is reserved, a noise amplitude increase may push the total signal power over the blue line, producing bit errors or even user offline events. When noise margin is reserved, the access device can tolerate a certain noise amplitude increase, allowing the total signal power to stay between the blue and red lines. In this way, the access device achieves higher line stability.

Application

The activated noise margin is associated with the target noise, and maximum and minimum noise margins configured for the access device. Specifically, the activated noise margin is close to the target noise margin, and within the range outlined by the maximum and minimum noise margins. A higher reserved noise margin means less power for carrying bits and a lower transmission rate.

Noise margins, including target, maximum, and minimum noise margins, apply in both upstream and downstream directions.

Target noise margin

- Target noise margin refers to the noise margin required for an access device to initialize with a BER of 10^{-7} or smaller. The target noise margin applies during line training and does not take effect after a line is trained. The line must be initialized with a BER of 10^{-7} or smaller. After line training is complete, users can query the actual noise margin of the line, which is close to the target noise margin.
- The target noise margin is reserved during normal data communication and it ensures normal communication in unfavorable line conditions. A larger noise margin means a less probability for the access device to encounter data transmission errors, a safer access device, but a lower maximum rate. For practical applications, configure a proper target noise margin based on line conditions.
- The access device establishes xDSL line connections and determines their rates according to the target noise margin. An over-high target noise margin may cause a decrease in the activated line rate, and an over-low target noise margin may cause an unstable line.

Maximum noise margin

- For a line in good conditions, if the activated noise margin exceeds the maximum noise margin, the access device must lower the line SNR by decreasing the signal power, while retaining the line rate up to the line requirement.
- In the process of xDSL connection establishment, if the noise margin calculated by the access device exceeds the specified maximum noise margin, the port will lower the signal power so that the noise margin will decrease to lower than the maximum noise margin.

Minimum noise margin

- When the line conditions turn unfavorable and the activated noise margin is lower than the minimum noise margin, the line cannot carry the expected bits. In this case, the line SNR must be raised by increasing the signal power so that the line can provide the required rate. If the signal power cannot be increased at all or cannot be increased to the extend to push the noise margin higher than the minimum noise margin, the line must be retrained.
- In the process of xDSL connection establishment, if the calculated noise margin is lower than the preset minimum noise margin, the port fails to be activated.

Determine the maximum and minimum noise margins based on line conditions. The maximum and minimum noise margin settings apply after the line is activated. A line keeps changing, sometimes in a good way and sometimes in a bad way.

- When the line condition worsens and the noise margin is lower than the minimum noise margin, the line cannot carry the expected bits. In this case, the line SNR must be raised by increasing the signal power so that the line can provide the required rate.
- When the line condition improves and the noise margin is higher than the maximum noise margin, the line SNR is over-high and will result in resource waste. In this case, the SNR must be lowered by decreasing the signal power, while the required line rate is retained.

An over-high target noise margin may decrease the activated rate, while an over-low target noise margin may result in an unstable line. Retain the default value (6 dB) for the target noise margin generally. If the activated rate is required at 0 km, the target noise margin can be reduced to a certain extent, but it is recommended that you retain the value greater than 3 dB; otherwise, the line may be unstable. In other conditions, the default value is recommended.

Bit Swapping

Bit swapping automatically adjusts the bit and power allocation on different tones according to SNR changes, so that the line is dynamically adaptive to variable noise without retrainings.

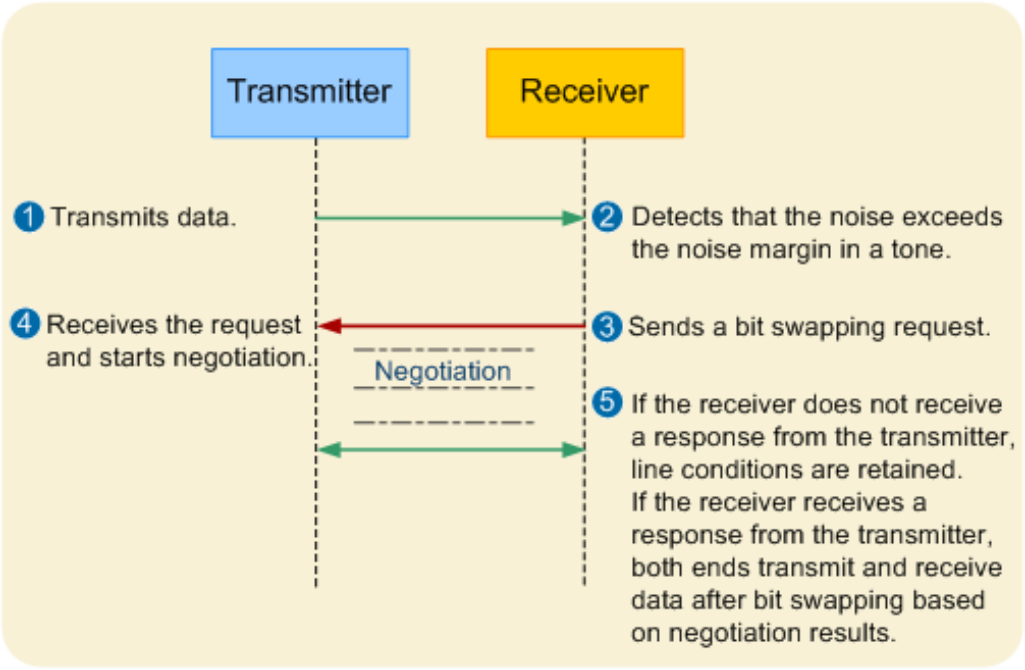
When the DSL line SNR changes but does not exceed the noise margin, the line BER meets the requirement (lower than 10^{-7}). However, noise margin does not always apply. When the line SNR decreases below the noise margin, the line BER will exceed 10^{-7}, and if it lasts for a long time, the line will be retrained to be adaptive to the noise.

Bit swapping automatically adjusts the bit and power allocation on different tones according to SNR changes, so that the line is dynamically adaptive to variable noise without retrainings.

As an online reconfiguration (OLR) technique, bit swapping does not change the line rate.

Figure 5-11 shows the working principle of bit swapping.

Figure 5-11 Working principle of bit swapping



- When detecting noise exceeding the noise margin in a tone, the receiver sends requests to the transmitter, requesting the transmitter to: swap bits from low-SNR tones to high-SNR tones; reduce the transmit power of the tones with reduced bits (crosstalk will result if these tones retain the original transmit power); increase the transmit power of the tones with increased bits.
- After the receiver sends bit swapping requests, the transmitter and receiver negotiate. Specifically, if the receiver does not receive response within a certain period of time, it deems that the transmitter does not support bit swapping (for example, when bit swapping is disabled) and retains the line conditions. If the receiver receives response from the transmitter, the transmitter and receiver will operate based on the negotiation results, to transmit or receive data. As devices (especially modems) supplied by different manufacturers have varied implementation of bit swapping, the transmitter and receiver, while negotiating and interacting with each other, may misunderstand each other. When misunderstanding happens, the line may be deactivated.

The Huawei access device allows users to enable or disable bit swapping in the upstream and downstream directions by running the **xdsl line-spectrum-profile add** command.

Seamless Rate Adaptation (SRA)

Bit swapping adjusts bit distribution on tones for a line to be noise-adaptive while retaining a constant rate. Seamless rate adaptation (SRA) enables the line to dynamically adapt to noises to a greater extent without retrainings.

When line conditions turn unfavorable and bit swapping fails to retain the bit error ratio (BER) at the required level, SRA decreases the rate; when line conditions turn favorable again, SRA increases the rate. In this manner, bandwidth usage is maximized.

Concepts

Association between line rates and bits

Line rate refers to the sum of bits transmitted over all tones on a channel.

SNR margin for rate upshift: When the noise margin reaches the specified value and sustains for **minimum upshift time**, the transmission rate automatically upshifts. **SNR margin for rate upshift** can be specified separately in upstream and downstream directions.

SNR margin for rate downshift: When the noise margin reaches the specified value and sustains for **minimum downshift time**, the transmission rate automatically downshifts. **SNR margin for rate downshift** can be specified separately in upstream and downstream directions.

Minimum upshift time: If the signal-to-noise ratio (SNR) margin reaches the value where the transmission rate starts to upshift, the transmission rate holds at this point for the specified minimum time and upshifts. **Minimum upshift time** can be specified separately in upstream and downstream directions.

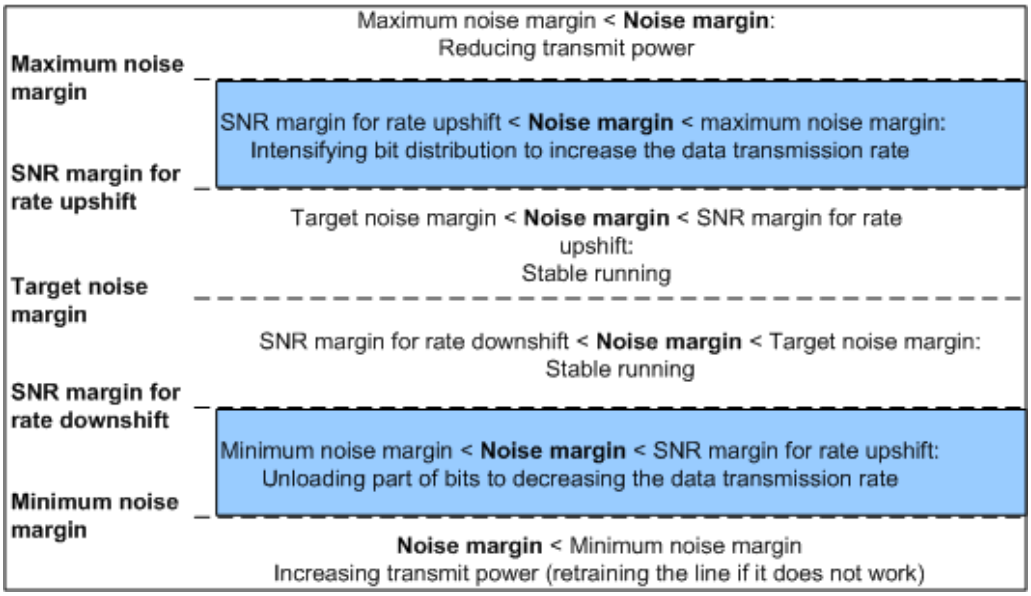
Minimum downshift time: If the SNR margin reaches the value where the transmission rate starts to downshift, the transmission rate holds at this point for the specified minimum time and downshifts. **Minimum downshift time** can be specified separately in upstream and downstream directions.

Working Principle

Figure 5-12 shows the association between a noise margin and SRA. The green-shaded blocks include description of SRA functions and the noise margin range.

- When **noise margin** is greater than or equal to **SNR margin for rate upshift** for over **minimum upshift time**, SRA functions to intensify bit distribution on the line for the transmission rate (line rate) to upshift.
- When **noise margin** is less than or equal to **SNR margin for rate downshift** for over **minimum downshift time**, SRA functions to unload part of bit distribution on the line for the transmission rate (line rate) to downshift.
- When **noise margin** is less than **SNR margin for rate upshift** but greater than **SNR margin for rate downshift**, or stays shorter than the minimum time, SRA will not function.

Figure 5-12 Noise margin

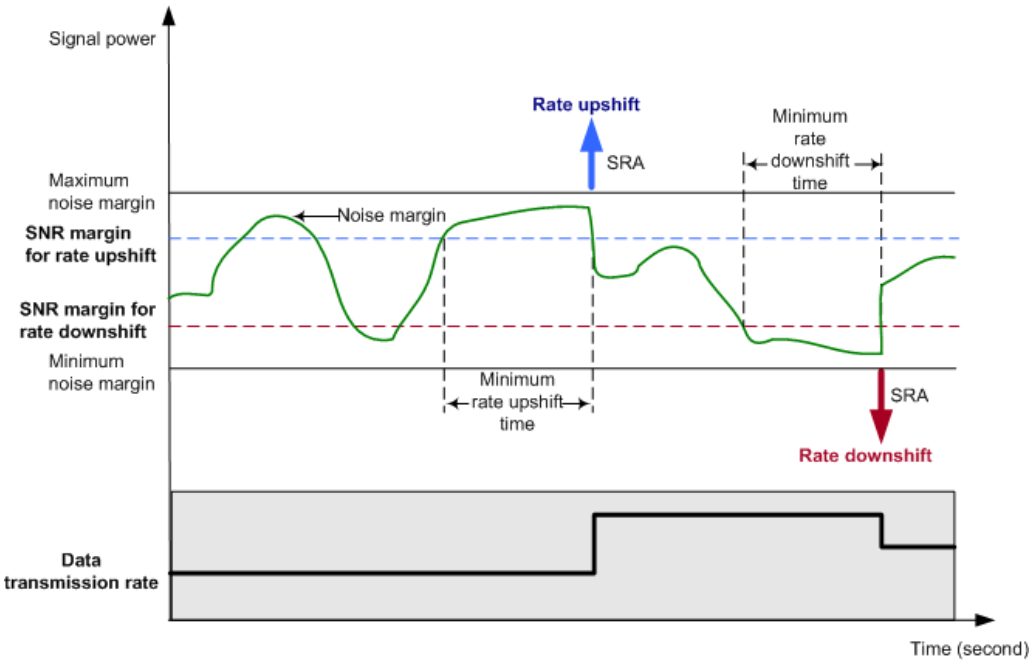


- When the noise margin is decreasing to lower than the SNR margin for rate downshift (which lies between the minimum and target noise margins), the customer premises equipment (CPE) sends control messages to the central office (CO), requesting the CO to dynamically decrease the signal transmit rate. After the signal transmit rate downshifts, the line noise margin increases. When the noise margin increases to the target value, the signal transmit rate stays stable.
- When the noise margin is increasing to higher than the SNR margin for rate upshift (which lies between the maximum and target noise margins), the CPE sends control messages to the CO, requesting the CO to dynamically increase the signal transmit rate. After the signal transmit rate upshifts, the line noise margin decreases. When the noise margin decreases to the target value, the signal transmit rate stays stable.

The rate upshift and downshift do not cause line retrainings or service interruption. This is why the rate adaptation process is seamless.

Figure 5-13 shows the entire SRA process and the specific process where the CO controls SRA using parameters.

Figure 5-13 SRA process



The rate does not upshift or downshift immediately when the line noise margin reaches the SNR margin for rate upshift or downshift. Instead, SRA starts to function only after the line noise margin stays at the level for the required time (in a range of 0s to 16383s).

Application

SRA can be enabled or disabled for an activated line. The receiver (CPE) triggers SRA while the transmitter (CO) controls SRA parameters.

SRA is sufficient to resolve the issues caused when noise margin changes slowly, but is insufficient when noise margin changes sharply.

Tone Blackout

If a certain band on the DSL line has unstable noise, which may cause interference, tone blackout can forbid the band from transmitting data, hence eliminating the interference. Some bands may be used for special purposes in certain regions; to prevent interference with these bands, tone blackout can forbid these bands.

Tone blackout, or missing tone as called in ADSL standards, means that a subcarrier is disabled and it will not carry any power (though there is a negligible transition band at both ends of the blackout band, because of the analog components), or any bit.

On the Huawei access device, users can run the **xdsl line-spectrum-profile add** command to configure tone blackout. The tone blackout band cannot be over-extensive or include the pilot tone; otherwise, the line may fail to be activated.

NOTE

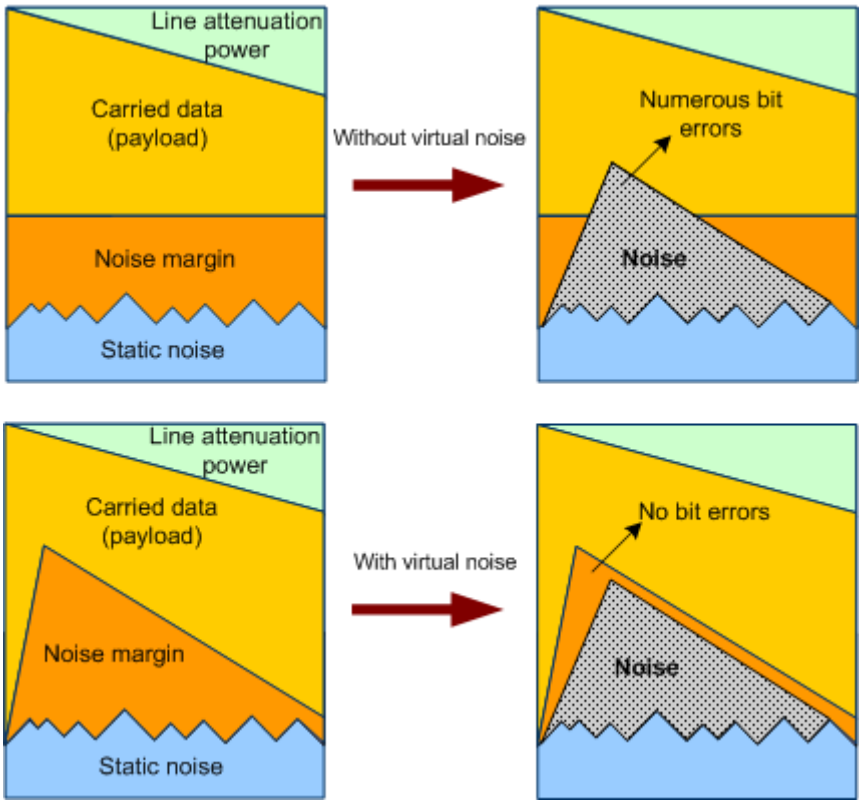
The system determines the pilot tones in line with ITU-T Recommendation G.994.1. Users can identify the pilot tones by comparing the spectrum profile against the ITU-T Recommendation G.994.1. Generally, the tone blackout band has a high frequency while the pilot tone has a low frequency, and they are less likely to intersect.

Virtual Noise

Noise margin is constant but line noise changes (the change fits a function of frequency). An over-large noise margin means fewer bits carried over tones and compromised performance; an over-small noise margin means a high BER when noise of a tone exceeds the noise margin. To resolve the issues, the noise margin power spectral density (PSD) mask must resemble the noise PSD mask whenever possible. This is how virtual noise helps.

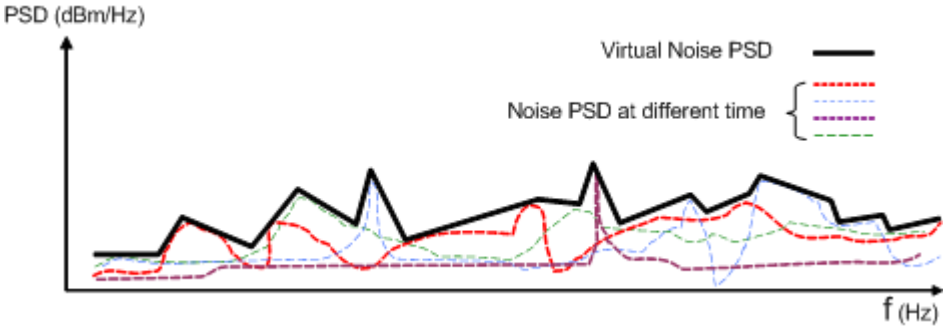
Figure 5-14 shows a reference model of virtual noise.

Figure 5-14 Reference model of virtual noise



For the virtual noise PSD mask to resemble the noise PSD mask in practical application, statistics on noise of the entire spectrum over a long period must be collected, as shown in Figure 5-15.

Figure 5-15 Virtual noise PSD mask



As shown in Figure 5-15, the virtual noise PSD mask more resembles the noise PSD mask than the noise margin, and ensures a more stable line and better line performance. In the meanwhile, however, virtual noise always presumes the maximum noise under the most unfavorable conditions. Therefore, line stability and low BER are achieved by compromising the connection rate.

NOTE

Figure 5-15 shows the statistical results as an example. In practical application, different carriers may use different tools and methods for collecting and analysing statistics, and the present of the statistical results may be different..

In line with ITU-T Recommendation G.997.1, the Huawei access device allows users to enable or disable virtual noise, and configure the noise margin profile and virtual noise profile by running the **xdsl noise-margin-profile add** and **xdsl virtual-noise-profile add** commands, respectively. A virtual noise profile includes multiple virtual noise PSD breakpoints. Based on this profile, the system draws the virtual noise mask for the entire spectrum using an interpolation algorithm. This process is similar to that for drawing a management information base (MIB) PSD mask.

5.3.2 Techniques for Reducing Interference

To minimize mutual interference between VDSL2 and other transmission systems, VDSL2 uses flexible mechanisms for controlling the transmit power. As these mechanisms shape the power spectral density (PSD), they are referenced as PSD shaping.

MIB-controlled PSD Mask

ITU-T Recommendation G.993.2 defines management information base (MIB)-controlled power spectral density (PSD) mask for a system to flexibly control PSD. "MIB-controlled" means configuring PSD masks through the network management system (NMS) or through a digital subscriber line access multiplexer (DSLAM). MIB-controlled PSD masks provide users with more options than the limit PSD masks defined in the standard. Carriers can control the power spectrum and reduce crosstalk by configuring suitable PSD masks according to DSLAM distribution, distance to users, and coexistence of ADSL and VDSL. Such user-configured PSD masks are referred to as MIB-controlled PSD masks.

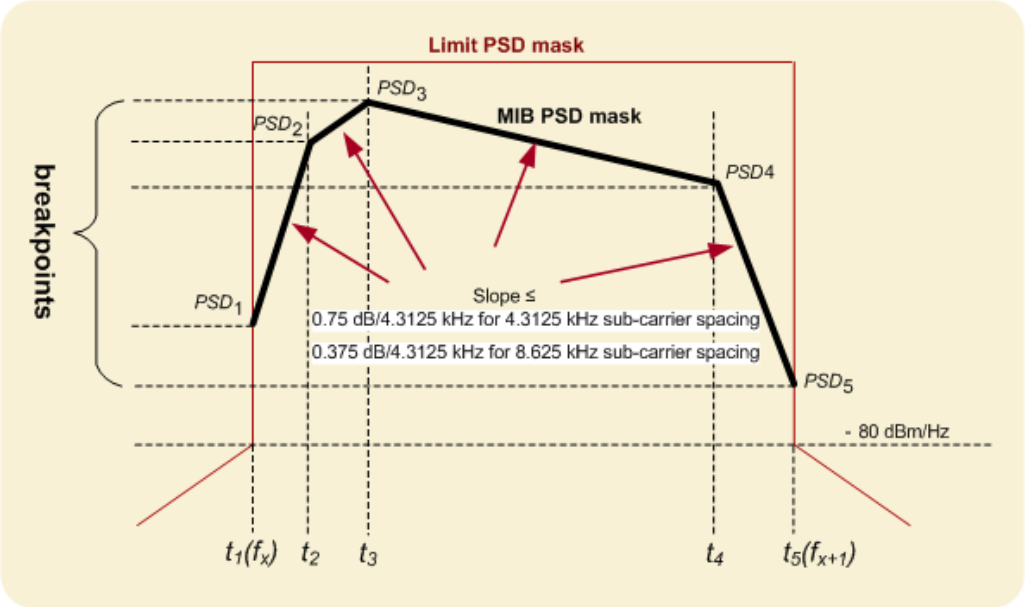
Figure 5-16 shows a common MIB-controlled PSD mask defined in ITU-T Recommendation G.993.2.

- The MIB-controlled PSD mask defines the PSD at a series of breakpoints on the transmission frequency band. Based on the PSD mask, the system determines the PSD of each subcarrier (or tone) using interpolation between two breakpoints.

- For each breakpoint, a subcarrier index (t_n) and PSD value (PSD_n) are defined. Then breakpoints are expressed like [(t_1, PSD_1), (t_2, PSD_2),…, (t_n, PSD_n)], where t_1 indicates the start frequency and t_n the stop frequency of the frequency band.
- In Figure 5-16, the limit PSD mask only indicates that the MIB-controlled PSD mask should always lie below the limit PSD mask (if the former lies above the latter, the system chooses the smaller one as the PSD mask). The turns at the PSD mask cannot form a right angle, and the slope for each turn is restricted to avoid a sharp change in the transmit power.

In addition, a maximum of 16 breakpoints can be configured in the upstream direction (for ADSL2+, a maximum of 4 breakpoints can be configured in the upstream directio) and 32 in the downstream direction. The US0 band cannot include any breakpoint.

Figure 5-16 MIB-controlled PSD mask



On the Huawei access device, users can configure MIB-controlled PSD masks by running the **xdsl mode-specific-psd-profile add** command.

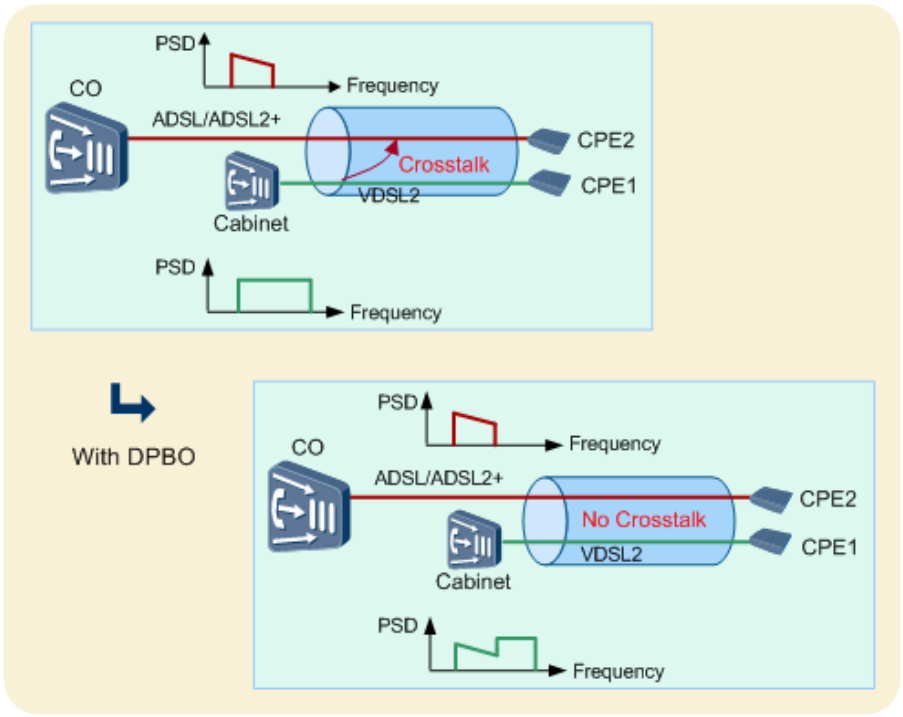
DPBO

Downstream power back-off (DPBO) is implemented to minimize crosstalk among the upstream lines in the same bundle (VDSL2 and ADSL/ADSL2+).

Definition of DPBO

On most conditions, VDSL2 lines are shorter than ADSL/ADSL2+ lines. This is why ADSL/ADSL2+ is deployed at CO and VDSL2 at cabinets, which are close to users, as shown in Figure 5-17.

Figure 5-17 Minimizing Inter-Line Crosstalk



Generally, after signals reach a cabinet, the downstream transmit power of CO is attenuated to far lower than the downstream transmit power of the cabinet. If VDSL2 and ADSL/ADSL2+ lines are deployed in the same cable bundle, the downstream signals of the cabinet have intensive crosstalk with the downstream signals of CO, which may be as intensive as to cause BER over $^{-7}$ and deteriorate services.

To minimize the inter-line crosstalk, DPBO is implemented to decrease the downstream transmit power of the cabinet so that it is close to the power of the CO-transmitted signals reaching the cabinet. Then the inter-line crosstalk is minimized.

ITU-T G.997.1 defines an algorithm for calculating DPBO, or the cabinet-end DPBO PSD mask. More specifically, the CO-end downstream PSD minus the power attenuated over the L (distance between the CO and cabinet) is equal to the PSD from the CO to cabinet. Then the cabinet-end downstream PSD is adjusted to close to the PSD.

DPBO Configuration

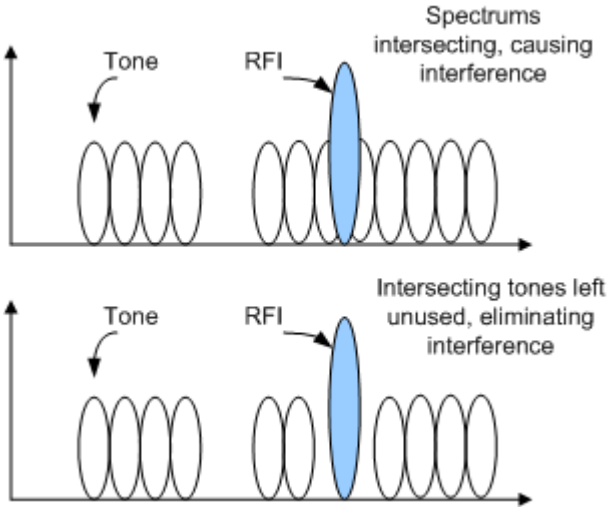
For DPBO to apply, some parameters regarding DPBO PSD mask calculation must be configured. For a Huawei access device, DPBO parameters include standard ones defined in ITU-T G.997.1, and non-standard ones customized for carriers (for ADSL2+, does not contain the non-standard ones). Users can configure DPBO by running the **xdsl dpbo-profile add** commands. For details on the parameters, see the description of the **xdsl dpbo-profile add** command.

PSD Notching

VDSL2 uses a wide range of frequencies, with the highest frequency of 30 MHz, which covers the medium wave, short wave, and ham radio. Therefore, VDSL2 has to provide a solution to radio frequency interference (RFI). There are complex RFI factors, and the

conventional countermeasures against RFI are not cost-effective. RFI lasts long and has such a narrow interference band that it is densely populated on one or several tones. RFI notching is introduced to resolve the issue.

Figure 5-18 Working principle of RFI notching



RFI notching means leaving some RFI-free tones unused to counteract RFI. Though RFI notching sacrifices some line transmission rate, it is effective. When the tones are left unused, the transmit PSD will be decreased to below the ITU-T Recommendation G.993.2-defined -80 dBm/Hz but not to none. If the tones can still carry bits with the transmit PSD below -80 dBm/Hz, the tones will carry some bits. This is how RFI notching differs from tone blackout.

In practical application, if the RFI power is intensive (no specific benchmark for the intensity), RFI notching may fail to eliminate RFI. In this case, tone blackout can black out the interference-suffering tones to avoid RFI.

On the Huawei access device, users can run the **xdsl rfi-profile add** command to configure RFI notching. The RFI notching band cannot be over-extensive or include the pilot tone; otherwise, the line may fail to be activated.

NOTE

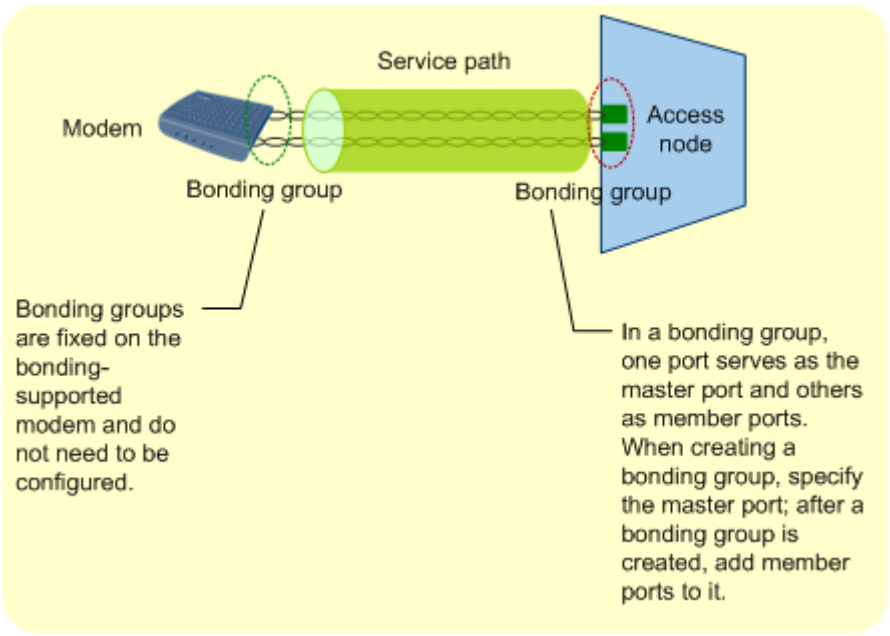
The system determines the pilot tones in line with ITU-T Recommendation G.994.1. Users can identify the pilot tones by comparing the spectrum profile against the ITU-T Recommendation G.994.1. Generally, the RFI notching band has a high frequency while the pilot tone has a low frequency, and they are less likely to intersect.

5.3.3 ADSL2+ ATM Bonding

ADSL2+ ATM bonding is implemented in line with ITU-T Recommendation G.998.1. It extends the access distance while maintaining a constant access rate or increases the access rate while maintaining a constant access distance, by means of bonding.

ADSL2+ ATM bonding supports bonding of two twisted pairs. Two ADSL2+ ports form a bonding group, one serving as master port and the other as member port, as shown in Figure 5-19. Services can be configured only on the master port in a bonding group.

Figure 5-19 Application of ADSL2+ ATM bonding

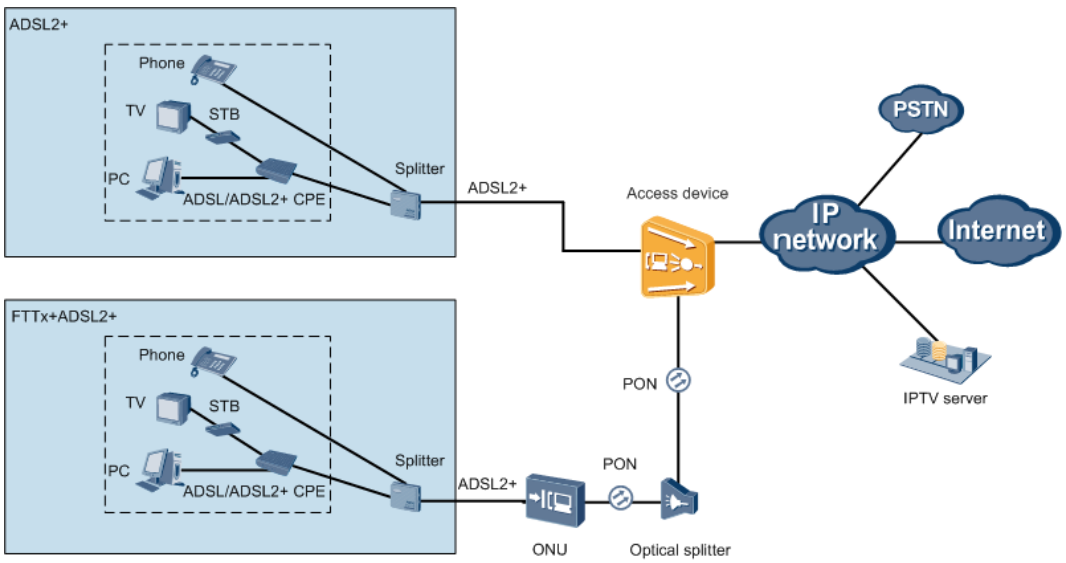


5.4 ADSL2+ Deployment and Maintenance

5.4.1 ADSL2+ Network Applications

This topic describes the network applications of the ADSL2+ access feature.

Figure 5-20 ADSL2+ network applications



As shown in the figure above, typical scenarios of ADSL2+ network applications are as follows.

1. The MA5600T/MA5603T/MA5608T provides the ADSL2+ access. On the user side, ADSL/ADSL2+ CPEs (working in the ATM mode) can be connected to the MA5600T/MA5603T/MA5608T to provide high-speed Internet access service, video service, and PSTN voice service for subscribers.
2. The MA5600T/MA5603T/MA5608T provides GPON optical ports for connecting to ONUs and the ONUs provide the ADSL2+ access. In the upstream direction, the ONUs are connected to the MA5600T/MA5603T/MA5608T by PON.

5.4.2 Brief Introduction to ADSL2+ Configurations and Applications

This topic provides a brief introduction to ADSL2+ configurations and applications. Such information helps you get an overview of ADSL2+ before moving on to details about its implementation.

ADSL2+ configuration includes two important steps:

1. Set spectrum parameters (for details, see 5.2.1 Spectrum Plan).
   a. Choose an appropriate transmission mode (that is, the applied standard and 6.3.2 Annex Types and US/DS Frequency Band Planning) depending on the DSL network plan and deployment.
   b. Configure 6.3.6 PSD Profiles based on power spectrum requirements. (You can manually configure a 6.3.9 MIB PSD Mask.)
2. Set line parameters to achieve a balance between performance and reliability (for details, see 6.4.2 Key Techniques for Improving Line Protection and 6.4.3 Techniques for Reducing Interference).

   Various noise interferences exist on a subscriber digital line. ADSL2+ provides a number of countermeasures to improve line stability, achieving higher line quality, and a lower packet loss ratio and bit error ratio. In most cases, stability is improved at the expense of line performance, for example, by reducing the activation rate or prolonging service latency. It is necessary, therefore, to set appropriate line parameters to achieve a balance between line reliability and performance. Table 5-3 lists the impact of various noise-cancellation countermeasures on line performance.

Table 5-2 Impact of countermeasures on line performance

Category	Countermeasure	Activation Rate Affected or Not	Service Latency Prolonged or Not
Improving line protection capabilities (passive defense against noise interference)	Interleaving FEC	Yes	Yes
	Configurable INP Parameters	Yes	Yes
	Physical Layer Retransmission (G.INP)	Yes	Yes
	Configurable Noise Margin	Yes	N/A
	Bit Swapping	N/A	N/A

Category	Countermeasure	Activation Rate Affected or Not	Service Latency Prolonged or Not
	SRA	No; the line rate is dynamically adjusted after a line is activated.	Yes (SRA may change the interleaving depth, resulting in latency deviations.)
	Tone Blackout	Yes	N/A
	Virtual Noise	Yes	N/A
Reducing interference output These countermeasures mitigate the impact of a line on other transmission systems. To achieve this, noise interference on the line must be reduced, mainly by means of power spectrum density (PSD) shaping	MIB-controlled PSD Mask	Yes	N/A
	DPBO	Yes	N/A

Table 5-3 lists techniques that counter different types of noises.

Table 5-3 Types of noises and countermeasures

Noise Type	Noise Characteristics	Countermeasure	Description
Pulse noises	Pulse noises are intensive, brief (micro- or milliseconds), and cover the entire frequency band. Pulse noise may derive from on-hook/off-hook of telephones,	• Interleaving FEC • Configurable INP Parameters • Physical Layer Retransmission (G.INP)	• Interleaving FEC, when used with erasure decoding, greatly improves system noise resistance. • To help users select appropriate INP

Noise Type	Noise Characteristics	Countermeasure	Description
	power-on/power-off of home appliances, or natural electricity discharge.		parameter values during configuration, ADSL2+ introduces the impulse noise monitoring (INM) technique. For details on erasure decoding and INM, see Configurable INP Parameters.
Environmental noises, such as background noise and noise caused by changes in temperature or relative humidity levels.	Noise that lasts a long period of time (microseconds), covers a narrow spectrum range, has a weak intensity, and changes slowly. Such a noise may come from amateur radio interference (such as that generated by remotely-controlled toys) and may overlap with radio frequency interference (RFI) described below.	Bit Swapping	In ITU-T Recommendation G.993.2, bit swapping, SRA, and SOS are on-line reconfiguration (OLR) techniques.
	Noise that lasts a long period of time (seconds), covers a wide spectrum range, has a weak intensity, and changes slowly. This type of noise shares some characteristics with inter-line crosstalk noises, as described below.	SRA	
	Noise that lasts a long period of time (seconds), covers a wide spectrum range, has a strong intensity, and changes fast. This type of noise shares some characteristics with inter-line crosstalk	SOS (seldom applied; not detailed here)	

Noise Type	Noise Characteristics	Countermeasure	Description
	noises, as described below.		
	Noise that lasts a long period of time (seconds), covers a wide spectrum range, and has a constant intensity. This type of noise shares some characteristics with inter-line crosstalk noises, as described below.	• Configurable Noise Margin • Virtual Noise	Currently, the Configurable Noise Margin technique is widely used.
RFI	RFI noise covers a narrow spectrum range, and interference occurs mostly on one or more tones. This type of noise mainly derives from broadcast and amateur radio communication.	• Tone Blackout • Bit Swapping	N/A
Inter-line crosstalk	Inter-line crosstalk refers to the noise caused by crosstalk between lines in a bundle, and it is associated with distribution of DSLAMs, distance to users, and coexistence of ADSL and VDSL2.	• DPBO • MIB-controlled PSD Mask • Bit Swapping • Configurable Noise Margin • Virtual Noise	The DPBO technique is recommended.

5.4.3 Configuration ADSL2+

ADSL2+ service configuration includes ADSL2+ profile configuration and ADSL2+ user port configuration. This topic describes the detailed configuration methods and procedures.

Overview of Configuring ADSL2+ Templates and Profiles

As mentioned in 5.4.2 Brief Introduction to ADSL2+ Configurations and Applications, spectrum parameter and line parameter configurations are the key points in ADSL2+ configuration. Spectrum and line parameters are configured in an ADSL2+ line parameter profile. In addition to the line parameter profile, the ADSL2+ line alarm profile can be configured to facilitate line maintenance. After the line parameter profile and line alarm

profile are configured, they can be directly used for activating DSL ports. The following describes the configuration of each ADSL2+ profile.

Context

The MA5600T/MA5603T/MA5608T supports three ADSL modes: normal (RFC2662), NGADSL (RFC4706), and TR165. Run the **switch adsl mode to** command to switch between the modes. By default, the normal mode is used.

- Normal mode: used for common ADSL2+ profiles, including ADSL2+ line profiles, line alarm profiles, and extended line profiles.
- NGADSL mode: ADSL2+ line profile parameters are reorganized, and a line template and a line alarm template are used. The line template uses the line profile and the channel profile, and the line alarm template uses the line alarm profile and the channel alarm profile.
- TR165 mode: A line profile consists of 10 profiles, which are: xDSL rate profile, power spectrum density (PSD) profile, xDSL spectrum profile, xDSL upstream power backoff (UPBO) profile, xDSL downstream power backoff (DPBO) profile, radio frequency interference (RFI) profile, xDSL noise margin profile, xDSL virtual noise profile, xDSL impulsive noise protection profile, and xDSL impulsive noise monitoring profile. All these profiles must be bound to an xDSL port for activating the xDSL port.

Configuring an ADSL2+ Alarm Template

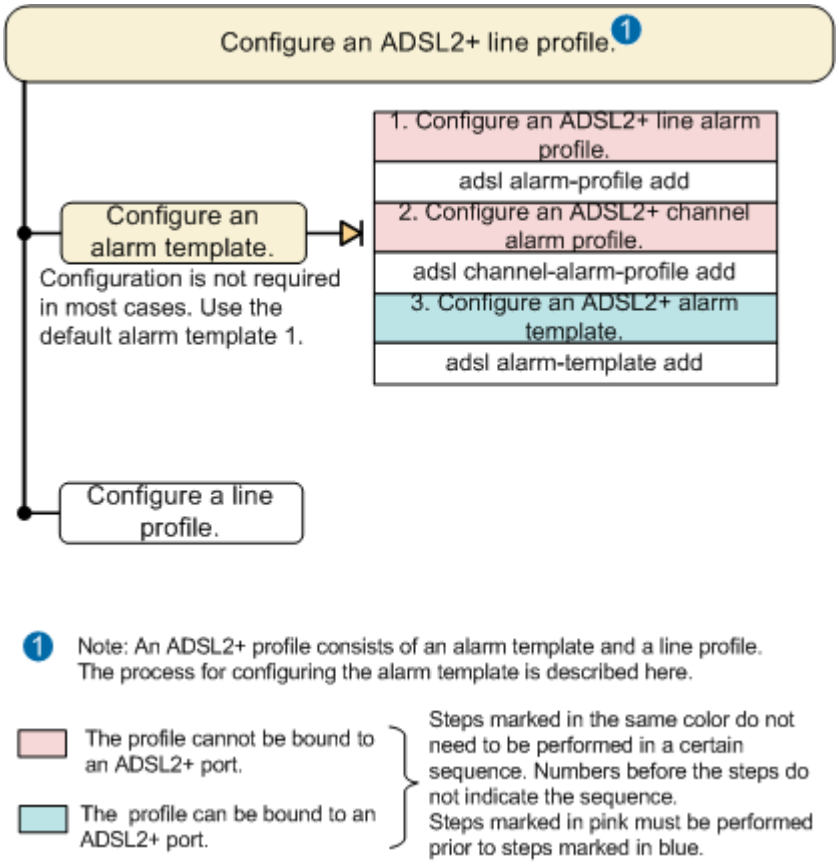
In RFC4706 and TR165 modes, an ADSL2+ alarm template that is used for activating ports consists of a line alarm profile and a channel alarm profile. The RFC2662 mode supports only line alarm profiles.

Context

In most cases, there is no need to configure an ADSL2+ alarm template. You can use the default alarm template 1.

If you want to configure the ADSL2+ alarm template, follow the process described in Figure 5-21.

Figure 5-21 Flowchart for configuring an ADSL2+ alarm template



Procedure

Configure an ADSL2+ line alarm profile.

Run the **adsl alarm-profile quickadd** command to quickly add an ADSL2+ line alarm profile, or run the interactive command **adsl alarm-profile add** to add an ADSL2+ line alarm profile.

Step 1 Configure an ADSL2+ channel alarm profile.

Run the **adsl channel-alarm-profile quickadd** command to quickly add an ADSL2+ channel alarm profile, or run the interactive command **adsl channel-alarm-profile add** to add an ADSL2+ channel alarm profile.

Step 2 Configure an ADSL2+ alarm template.

Run the **adsl alarm-template quickadd** command to quickly add an ADSL2+ alarm template, or run the interactive command **adsl alarm-template add** to add an ADSL2+ alarm template.

The main parameters are as follows:

- **line** *alarm-profile-index*: indicates the line alarm profile in the alarm template. If this parameter is required, configure it prior to **channel1**.
- **channel1** *channel1-alarm-profile-index*: indicates the channel alarm profile for channel 1 in the alarm template.

- **channel1** *channel1-alarm-profile-index*: indicates the channel alarm profile for channel 2 in the alarm template. Channel 2 is unavailable and this configuration will not take effect. Therefore, there is no need to set this parameter.

Step 3 Verify that the configurations in the alarm template agree with the data plan.

Run the **display adsl alarm-template** command to check whether the configurations in the alarm profile agree with the data plan.

After the alarm profile is successfully configured, it can be directly used for activating ADSL2+ ports.

----End

Example

The following configurations are used as an example to add alarm template 3:

- Alarm template 3 uses line alarm profile 2 and channel alarm profile 1 (default).
- The function of reporting terminal power-off alarms is disabled in line alarm profile 2.

```
huawei(config)#adsl alarm-profile quickadd 2 dying-gasp-switch disable
huawei(config)#adsl alarm-template quickadd 3 line 2 channel1 1
huawei(config)#display adsl alarm-template 3
```

Configuring an ADSL2+ Line Profile

An ADSL2+ line profile is the key for ADSL2+ service configurations. This topic describes how to configure the ADSL2+ line profiles in different ADSL2+ modes.

Prerequisites

Run the **display xdsl mode** command to check whether the ADSL2+ mode is the desired mode. The default mode is RFC2662.

If the current mode is not the desired one, run the **switch adsl mode to** command in diagnose mode to switch the mode to the desired mode.

NOTE
When both the ADSL2+ and VDSL2 modes are TR165, the configured profile is used by both ADSL2+ and VDSL2 ports. If only one of the ADSL2+ and VDSL2 modes is TR165, the configured profile is used only by the one in TR165 mode.

Configuration Process

Figure 5-22 shows the process for configuring an ADSL2+ line profile in RFC2662 mode.

Figure 5-22 Flowchart for configuring an ADSL2+ line profile - RFC2662 mode

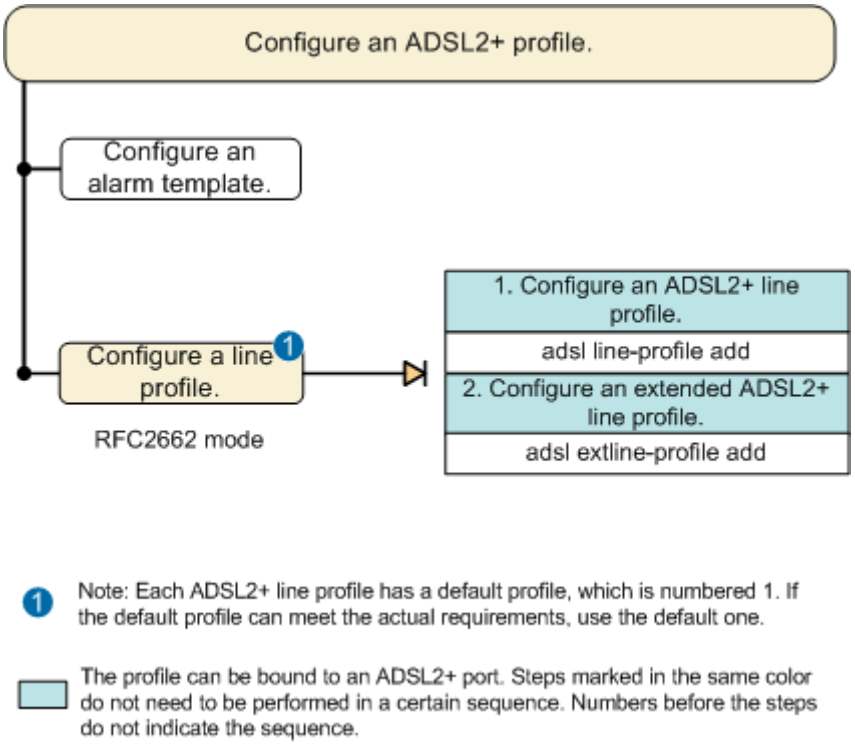


Figure 5-23 shows the process for configuring an ADSL2+ line profile in RFC4706 mode.

Figure 5-23 Flowchart for configuring an ADSL2+ line profile - RFC4706 mode

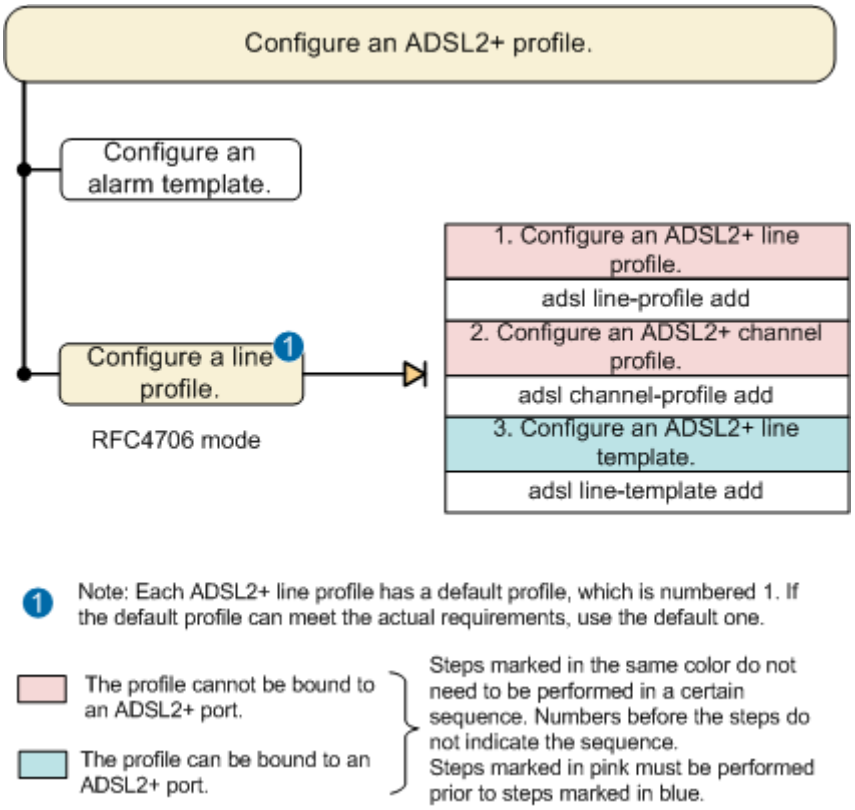
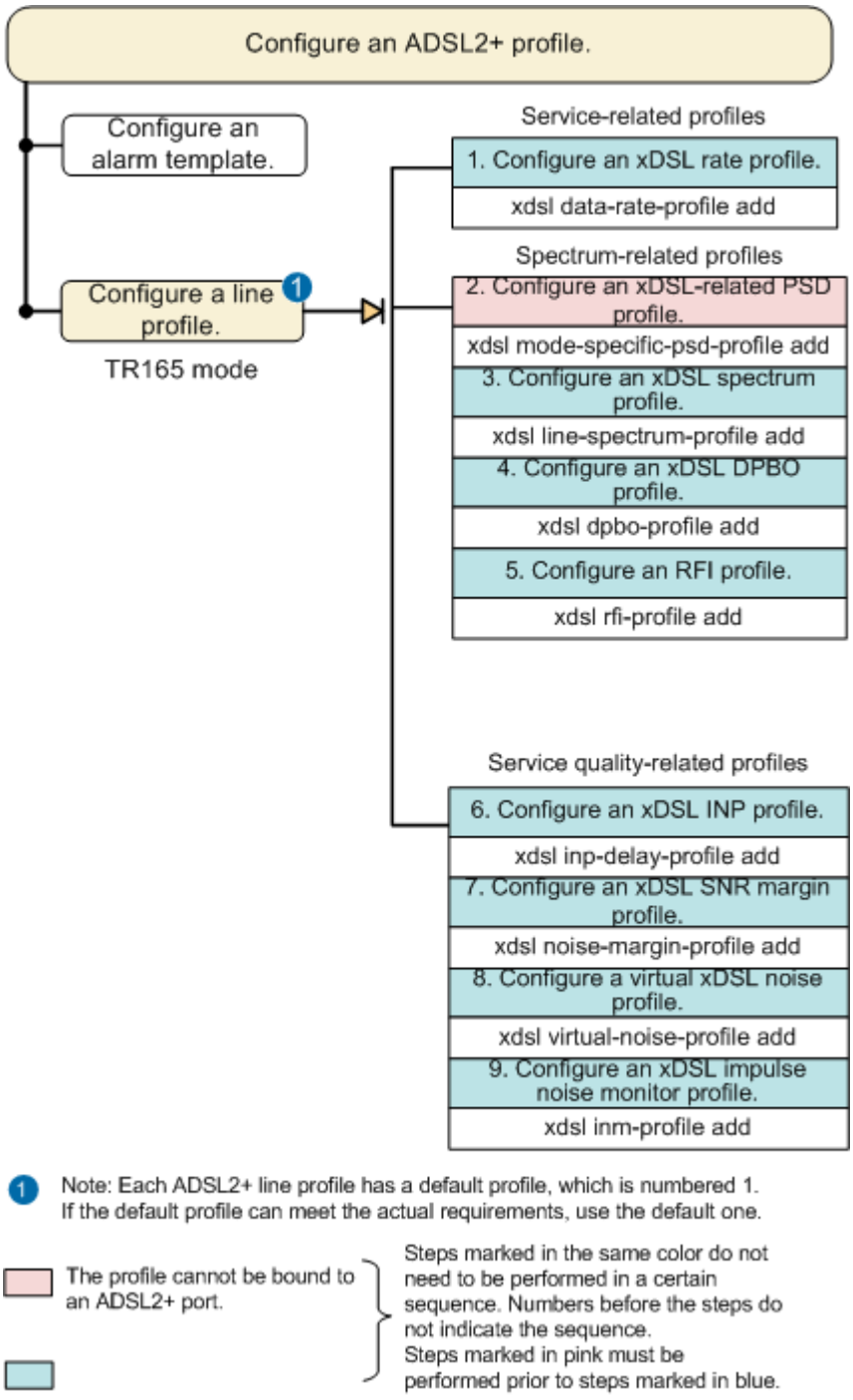


Figure 5-24 shows the process for configuring an ADSL2+ line profile in TR165 mode.

Figure 5-24 Flowchart for configuring an ADSL2+ line profile - TR165 mode



Procedure

- Do as follows to configure an ADSL2+ line profile when the ADSL2+ mode is RFC2662:
  a. Configure an ADSL2+ line profile.

Run the **adsl line-profile quickadd** command to quickly add an ADSL2+ line profile, or run the interactive command **adsl line-profile add** to add an ADSL2+ line profile.

An ADSL2+ port is bound to the ADSL2+ line profile when it is activated by running the **activate** command.

b. Configure an extended ADSL2+ line profile.

Run the **adsl extline-profile quickadd** command to quickly add an extended ADSL2+ line profile, or run the interactive command **adsl extline-profile add** to add an extended ADSL2+ line profile.

Run the **extline-config** command to bind the extended ADSL2+ line profile to an ADSL2+ port.

- Do as follows to configure an ADSL2+ line profile when the ADSL2+ mode is RFC4706:

a. Configure an ADSL2+ line profile.

Run the **adsl line-profile quickadd** command to quickly add an ADSL2+ line profile, or run the interactive command **adsl line-profile add** to add an ADSL2+ line profile.

b. Configure an ADSL2+ channel profile.

Run the **adsl channel-profile quickadd** command to quickly add an ADSL2+ channel profile, or run the interactive command **adsl channel-profile add** to add an ADSL2+ channel profile.

c. Configure an ADSL2+ line template.

Run the **adsl line-template quickadd** command to quickly add an ADSL2+ line template, or run the interactive command **adsl line-template add** to add an ADSL2+ line template.

A line template binds a line profile to a channel profile. Therefore, when activating an ADSL2+ port, you only need to bind the line template to the ADSL2+ port.

After a profile is successfully configured, it can be used for activating ADSL2+ ports.

- Do as follows to configure an ADSL2+ line profile when the ADSL2+ mode is TR165:

a. Configure service-related profiles.

i. Configure an xDSL rate profile.

Run the **xdsl data-rate-profile quickadd** command to quickly add an xDSL rate profile, or run the interactive command **xdsl data-rate-profile add** to add an xDSL rate profile.

NOTE

- When ADSL2+ ports are activated in TR165 mode, the upstream rate profile and downstream rate profile are used separately. The two profiles can be one profile. However, they are usually two different profiles because the upstream and downstream rates are different in actual practice.
- It is recommended that the **Data path mode** parameter in this command take the default value. If this parameter does not take the default value, ensure that it has the same value in the upstream and downstream rate profiles that are used for activating an ADSL2+ port.

b. Configure spectrum-related profiles.

i. Configure an xDSL-related PSD profile.

Run the **xdsl mode-specific-psd-profile quickadd** command to quickly add an xDSL-related PSD profile, or run the interactive command **xdsl mode-specific-psd-profile add** to add an xDSL-related PSD profile.

ii. Configure an xDSL spectrum profile.

Run the **xdsl line-spectrum-profile quickadd** command to quickly add an xDSL spectrum profile, or run the interactive command **xdsl line-spectrum-profile add** to add an xDSL spectrum profile.

iii. Configure an xDSL DPBO profile.

Run the **xdsl dpbo-profile quickadd** command to quickly add an xDSL DPBO profile, or run the interactive command **xdsl dpbo-profile add** to add an xDSL DPBO profile.

iv. Configure an RFI profile.

Run the **xdsl rfi-profile quickadd** command to quickly add an RFI profile, or run the interactive command **xdsl rfi-profile add** to add an RFI profile.

When spectrum-related profiles (except mode specific PSD profiles) are successfully configured, they can be used for activating ADSL2+ ports. Mode specific PSD profiles are not directly used for activating ports but are used in spectrum-related profiles.

c. Configure service quality-related profiles.

i. Configure an xDSL INP profile.

Run the **xdsl inp-delay-profile quickadd** command to quickly add an xDSL INP profile, or run the interactive command **xdsl inp-delay-profile add** to add an xDSL INP profile.

ii. Configure an xDSL SNR margin profile.

Run the **xdsl noise-margin-profile quickadd** command to quickly add an xDSL SNR margin profile, or run the interactive command **xdsl noise-margin-profile add** to add an xDSL SNR margin profile.

iii. Configure a virtual xDSL noise profile.

Run the **xdsl virtual-noise-profile quickadd** command to quickly add a virtual xDSL noise profile, or run the interactive command **xdsl virtual-noise-profile add** to add a virtual xDSL noise profile

iv. Configure an xDSL impulse noise monitor profile.

Run the **xdsl inm-profile quickadd** command to quickly add an xDSL impulse noise monitor profile, or run the interactive command **xdsl inm-profile add** to add an xDSL impulse noise monitor profile.

NOTE

Users can determine the INP value based on the obtained INMAINPEQi and INMAIATi histogram to protect the line stability.

- **INM inter arrival time offset**: indicates the INM inter-arrival time offset (INMIATO). It determines the INMAIATi histogram parameter range with INMIATS. It also determines the start point of IAT.
- **INM inter arrival time step**: indicates the INM inter-arrival time step (INMIATS). It determines the INMAIATi histogram parameter range with INMIATO. It also determines the precision of IAT.
- **INM cluster continuation value**: indicates the INM cluster continuation (INMCC) value. It identifies a cluster and indicates the maximum number of consecutive undamaged DMT symbols allowed in a cluster.
- **INM equivalent INP mode**: Indicates the INM equivalent impulse noise protection (INP) mode. The method of calculating the equivalent INP varies according to the mode. Mode 3 is recommended because the algorithm for the mode is better than the algorithms for modes 0, 1, and 2.

After service quality-related profiles are successfully configured, they can be used for activating ADSL2+ ports.

----End

Example

The following configurations are used as an example to create ADSL2+ line template 3:

- Downstream rate: 2048 Kbits/s
- Channel mode: interleave
- Maximum interleave delay: 10 ms
- SNR margin: 6 dB

```
huawei(config)#adsl line-profile quickadd 3 snr 60 30 120 60 30 120
huawei(config)#adsl channel-profile quickadd 3 interleaved-delay 10 10 rate 1024
 2048 3096 1024 2048 3096
huawei(config)#adsl line-template quickadd 3 line 3 channel1 3 60 70 channel2 3
```

Configuring ADSL2+ Line Bonding

To ensure longer access distance at the same access rate or higher access rate in the same access distance, configure ADSL2+ line bonding.

Prerequisites

- The port to be bound has no service flow.
- The port to be bound is in the activating or deactivated state.

NOTE

- An xDSL port can be in any of the following states: activating, activated, deactivated, and loopback.
- For the H802ADPD board, you need to run the **board workmode bonding** to set the BONDING mode as the working mode. In the normal state, running this command successfully causes the board to reset. Therefore, exercise caution when running this command!

Procedure

Create a bonding group.

In global config mode, run the **bonding-group add** command to create a bonding group.

Key parameters:

- **primary-port**: indicates the primary port in the bonding group. After a bonding group is created, service flows can be created only on the primary port.
- **scheme**: indicates the local bonding mode, which can be **ATM**, **EFM**, or **TDIM**. ADSL2+ ports support ATM-based bonding and the local bonding mode must be set to **ATM**. That is, two ADSL2+ ports on the same board are added to a bonding group. Operations for an ADSL2+ bonding group are performed on the primary port.
- **peer-scheme**: indicates the peer bonding mode, which must be the same as **scheme**.

Step 1 Add member ports for a bonding group.

Run the **bonding-group link add** command to add member ports.

Step 2 (Optional) Create a bonding group profile and configure line parameters for the ports in the bonding group.

Run the **xdsl bonding-group-profile add** command to create a bonding group profile and set line parameters for ports in the bonding group.

- There is a default profile: profile 1.
- The priority of the bonding group profile is higher than the line parameter profiles of the ports in the bonding group. When both the bonding group profile and line parameter profiles of the ports are used, the bonding group profile takes effect. If the maximum and minimum upstream/downstream transmission rates are set to 0, the rates are not limited in the bonding group profile and are determined by the rate limits specified in the line parameter profiles of the ports.

Step 3 Activate a bonding group.

Run the **active bonding-group** command to activate a bonding group.

Step 4 Query information about a bonding group.

Run the **display bonding-group** command to query information about a bonding group.

----End

Example

The following configurations are used as an example to add bonding group 1:

- ADSL2+ ports 0/2/0 and 0/2/1 are added to bonding group 1.
- 0/2 is the primary port.
- Bonding group 1 is activated using bonding group profile 1 (default).

```
huawei(config)#bonding-group add 1 primary-port 0/2/0 scheme atm peer-scheme atm
huawei(config)#bonding-group link add 1 0/2/1
huawei(config)#active bonding-group 1 profile-index 1
```

Configuring ADSL2+ User Ports

xDSL ports must be activated before they are used to transmit services. This topic describes how to activate ADSL2+ ports and enables the ports to use ADSL2+ profiles.

Prerequisites

5.4.3 Configuration ADSL2+ has been completed based on the data plan.

Procedure

- Do as follows to configure the ADSL2+ user ports when the ADSL2+ mode is RFC2662:
  a. In global config mode, run the **interface adsl** command to enter the ADSL mode.
  b. Run the **deactivate** command to deactivate ADSL2+ ports.
  c. Run the **activate** command to activate ADSL2+ ports and enable them to use the ADSL2+ line parameter profiles.
  d. Run the **alarm-config** command to enable the ADSL2+ ports to use the ADSL2+ alarm template.
- Do as follows to configure the ADSL2+ user ports when the ADSL2+ mode is RFC4706:
  a. In global config mode, run the **interface adsl** command to enter the ADSL mode.
  b. Run the **deactivate** command to deactivate ADSL2+ ports.

c. Run the **activate** command to activate ADSL2+ ports and enable them to use ADSL2+ line template.

d. Run the **alarm-config** command to enable the ADSL2+ ports to use the ADSL2+ alarm template.

- Do as follows to configure the ADSL2+ user ports when the ADSL2+ mode is TR165:

a. In global config mode, run the **interface adsl** command to enter the ADSL mode.

b. Run the **deactivate** command to deactivate ADSL2+ ports.

c. Run the **activate** command to activate ADSL2+ ports and enable them to use ADSL2+ line parameter profiles.

d. Run the **alarm-config** command to enable the ADSL2+ ports to use the ADSL2+ alarm template.

----End

Example

The following configurations are used as an example to activate ADSL2+ port 0/2/0 in RFC2662 mode and enable the port to use ADSL2+ alarm template 3 and ADSL2+ line template 6:

```
huawei(config)#interface adsl 0/2
huawei(config-if-adsl-0/2)#deactivate 0
huawei(config-if-adsl-0/2)#activate 0 template-index 6
huawei(config-if-adsl-0/2)#alarm-config 0 3
```

5.4.4 ADSL2+ Maintenance and Fault Diagnosis

There are many maintenance and fault diagnosis methods for DSL lines. The following describes the common faults and troubleshooting methods.

Common ADSL2+ Line Faults and Troubleshooting Methods

The diagnosis and troubleshooting methods for common ADSL2+ line faults are described to facilitate line maintenance.

Common Faults on ADSL2+ Lines

1. When the line is activated for the first time,

- The line fails to be activated.
- The activation rate is slow.

2. When the line is normal operation, the line quality degrades and consequently the line rate decreases or even the line is deactivated.

Alarms and events involved in these faults are as follows:

- 0x29100001 The ring topology in the subscriber port is found
- 0x3d300007 The xDSL channel downstream rate is lower than the threshold
- 0x3d30000b The xDSL channel upstream rate is lower than the threshold
- 0x0a11a055 The ADSL port activation rate fails to reach the rate threshold

- 0x0a300013 The ADSL port is automatically deactivated due to loss of signal(LOS) or loss of frame(LOF)
- 0x0a300017 The performance statistics of the ADSL port reach the threshold
- 0x3d30000d The line performance statistics of the ADSL port reach the threshold
- 0x3d30000e The ADSL channel downstream activation rate is lower than the threshold

Causes of the Common Faults

Table 5-4 Causes of the common ADSL2+ line faults

Reason	Description	Troubleshooting
Line parameters are improperly configured.	The target SNR margin is improperly configured. A large margin may decrease the activation rate and a small margin may affect the stability of the line.	1. In ADSL mode, run the **display line operation** command to check if the value of **Line SNR margin downstream/upstream** is proper compared with the historical values or the value of a functional port. If the value is improper, follow instructions provided in Configuring an ADSL2+ Line Profile to modify SNR Margin configurations. Then reactivate the port using the new profile. 2. In global config mode, run the **display event history** command to check if the related events have been generated. If yes, clear the event by referring to the Alarm and Event Handling.
	The minimum INP is improperly configured. There is a restrictive relationship between INP and line activation rate. Under a certain interleave depth, the line activation rate decreases with the increase of the INP value. If the minimum INP is large (for example, 16), the maximum interleave delay must also be large (for example, 63 ms). If the minimum INP is large while the maximum interleave delay is small, the line activation rate will be low or even the activation fails.	1. In ADSL mode, run the **display parameter** command to check if the values of **Minimum impulse noise protection downstream/upstream** and **Maximum interleaving delay downstream/upstream** are proper. If the values are improper, follow the instructions provided in Configuring an ADSL2+ Line Profile to modify the configurations of the minimum INP and maximum interleave delay. Then reactivate the port using the new profile. 2. In global config mode, run the **display event history** command to check if the related events have been generated. If yes, clear the event by referring to the Alarm and Event Handling.
Physical lines are of poor	There are engineering issues. For example, the physical line is not securely connected or	1. Securely connect physical lines or replace the lines. 2. In global config mode, run the **display event history** command to check if the related events

Reason	Description	Troubleshooting
quality.	deteriorates.	have been generated. If yes, clear the event by referring to the Alarm and Event Handling.
	There is a loop in subscriber lines.	In global config mode, run the **display event history 0x29110001** command to check if a loop alarm has been generated. If yes, communicate with the subscriber that owns the alarming port and help the subscriber check its line connections and release the loop.
	There are interference sources around DSL lines.	Check if there are strong interference sources around subscriber lines, such as a wireless base station and high-frequency switch-mode power supply. 1. Remove the interference sources as much as possible or reroute the subscriber lines. 2. You can also deal with the interference by RFI Notching, Tone Blackout, increasing SNR margin, or limiting the activation rate.
	The ADSL2+ board or port is faulty.	Rectify the fault by referring to Loopback on an ADSL2+ Port.

Loopback on an ADSL2+ Port

A loopback on an ADSL2+ port can be performed to determine whether the service board of the ADSL2+ port is communicating with the backplane properly and accordingly locate the fault.

Prerequisites

- The service board running the ADSL2+ service is functioning properly.
- The ADSL2+ port is deactivated.
- The ADSL2+ service ran properly before the fault occurred. This ensures that a downstream service flow exists between the control board and the ADSL2+ service board.

Impact on the System

- A port involved in a loopback cannot forward data packets and all services carried on the port are interrupted.
- If a port involved in a loopback is not isolated, a broadcast storm may occur on the device of the port, which affects the services carried on other ports.

Before starting a loopback test on a port, set the test duration. After the loopback test is complete, run the **undo loopback** command to cancel the loopback.

Notes

- Connect a 100-ohm resistor in series to an ADSL2+ line to perform a hybrid loopback.

- The loopback type supported by a board is based on the hardware structure of the board. If a loopback is performed on a board that does not support the loopback, the MA5600T/MA5603T/MA5608T of the board displays an error message.
- Only one type of loopback can be performed on an ADSL2+ port at a time.

Procedure

In ADSL mode, run the **loopback** command to start a loopback.

NOTE
Port loopback is classified as local loopback and remote loopback. For details about local loopback and remote loopback, see section *Reference* in the following section.

The following configuration is used as an example to perform a loopback of the UTOPIA type on port 0/1/0:

```
huawei(config-if-adsl-0/1)#loopback 0 UTOPIA
```

NOTE
- A loopback can be of the universal test and operations PHY interface for ATM (UTOPIA), analog front end (AFE), or hybrid type.
- To start another loopback, run the **undo loopback** command to cancel the ongoing loopback.

Step 1 In ADSL mode, run the **atm-ping** command to check the connectivity of the loop path.

The following configurations are used as an example to perform a test for the preceding loop path with VPI/VCI 0/35 (VPI is the abbreviated form of virtual path identifier and VCI is that of virtual channel identifier):

```
huawei(config-if-adsl-0/1)#atm-ping 0 0 35
  atm-ping successfully. Sequence=0
  atm-ping successfully. Sequence=1
  atm-ping successfully. Sequence=2
  atm-ping successfully. Sequence=3
  atm-ping successfully. Sequence=4

  --- Atm-ping adsl0/1/0 0/35 statistics ---
  5 oam f5 loopback cells transmitted
  5 oam f5 loopback cells received
  0.00% cell loss
```

NOTE
- If the ping operation is successful and no packets are lost, the loop path is functional.
- If the ping operation fails, the loop path is disconnected.
- If the ping operation is successful but some packets are lost, the loop path is faulty.

Step 2 Run the **undo loopback** command to cancel the loopback after the loopback ends.

NOTE
A port on which a loopback is being performed cannot be activated.

----End

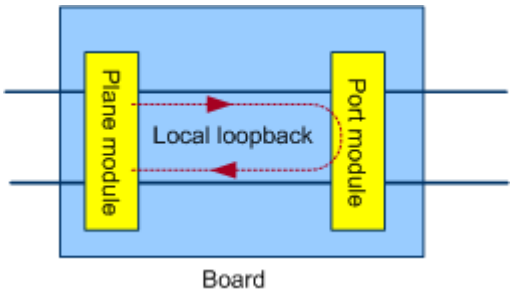
Reference

An ADSL2+ board supports only local loopbacks.

Introduction to a local loopback

Local loopback, also called inloop, near-end loopback, or central office (CO) loopback, is performed from the port processing module of a service board to the backplane. In this loopback, signals sent from the backplane to the port are returned to the backplane to check whether the service board is working properly. Figure 5-25 shows a local loopback.

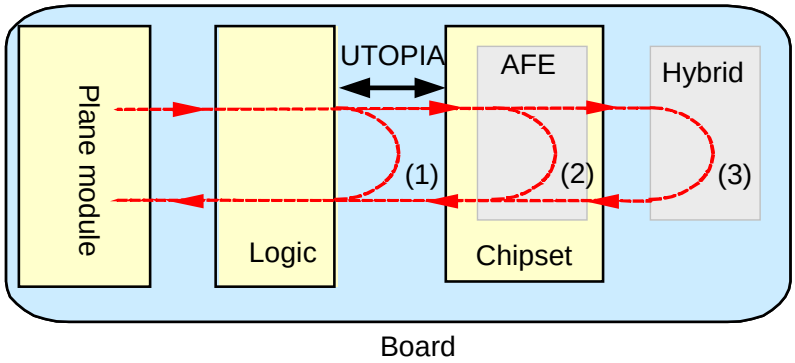
Figure 5-25 Local loopback



Local loopback on an ADSL2+ port

Figure 5-26 shows a local loopback on an ADSL2+ port.

Figure 5-26 Local loopback on an ADSL2+ port



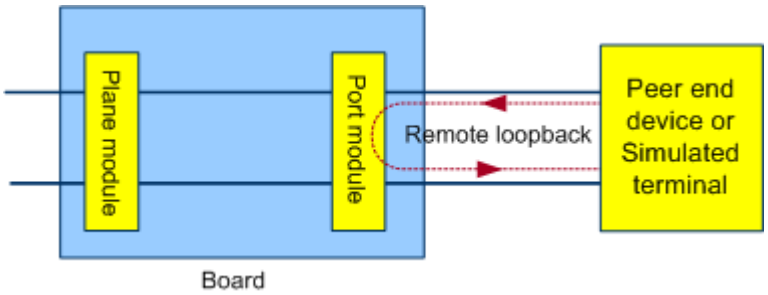
- In a UTOPIA loopback, signals are sent from the backplane to the UTOPIA interface and back to the backplane, as (1) in Figure 5-26 shows. This loopback checks whether the loop path between the backplane and the logic chip is functional.
- In an AFE loopback, signals are sent from the backplane to the AFE and back to the backplane, as (2) in Figure 5-26 shows. This loopback checks whether the loop path between the backplane and the chipset is functional.
- In a hybrid loopback, signals are sent from the backplane to the hybrid interface and back to the backplane, as (3) in Figure 5-26 shows. This loopback checks whether the loop path between the backplane and the chipset edge is functional.

Remote Loopback

Remote loopback, also called outloop, refers to the loopback from the port processing module inside the board to the subscriber line. In remote loopback, the signals between the user-side device (such as the modem) and the port signal receiving module directly return to the user-side device through the port signal sending module over the subscriber line. The test aims to check whether the upstream service between the customer premises equipment (CPE)

and the board is through, and whether packet loss exists. When the service failure occurs, the fault is located on the CPE or the board chip set. Figure 5-27 shows the remote loopback.

Figure 5-27 Remote loopback



5.5 Standard and Protocol Compliance

Table 5-5 lists the standards and protocols that ADSL2+ complies with.

Table 5-5 Standard and Protocol Compliance

Standard and Protocol	Description
ITU-T G.992.1	ADSL transceivers
ITU-T G.992.3	ADSL2
ITU-T G.992.5	ADSL transceivers – Extended bandwidth ADSL2 (ADSL2plus)
ITU-T G.997.1	Physical layer management for DSL transceivers
ITU-T G.998.1	ATM-based multi-pair bonding
ITU-T G.998.4	Improved impulse noise protection (INP) for DSL transceivers
Broadband Forum TR-159	Management framework for xDSL bonding

5.6 Appendix 1: Introduction to the ADSL2+ Coding/Decoding Technologies

ADSL2+ coding/decoding is essential for improving line quality and performance.

DMT Modulation

DMT divides transmission bandwidth into n stand-alone or discrete sub-carriers (also called tones) and performs orthogonal transforming on data segments in each sub-carrier. The most common transforming method is discrete Fourier transform (DFT). The data rate of each sub-carrier is 1/n of the entire data rate.

Pilot Tone

DMT requires strict clock synchronization between devices at both ends. For clock synchronization, several pilot tones can be inserted to avoid wandering of frequency points.

Optional Cyclic Extension Length

DMT supports a cyclic extension between DMT symbols and uses the cyclic extension for protection. This cyclic extension is also called cyclic prefix. A cyclic prefix eliminates the interference caused by latency extension between DMT symbols but lowers the bandwidth usage.

ITU-T Recommendation G.993.2 stipulates calculation of optional cyclic extension length. Specifically, if the path conditions are unfavorable, the cyclic prefix can be extended to prolong the protection interval, which helps eliminate interference between DMT symbols. If the path conditions are favorable, the cyclic prefix can be narrowed to increase bandwidth usage.

The Huawei access device enables users to run commands to set **Optional Cyclic Extension Flag** (enabled or disabled), which complies with ITU-T Recommendation G.997.1. **Optional Cyclic Extension Flag** identifies whether to enable the optional cyclic extension. If it is enabled, the algorithm for calculating the optional cyclic prefix is started; if it is disabled, the cyclic prefix of a fixed length is used.

Scrambling

Data transmitted over the line may contain long strings of consecutive 0s or 1s. Such data may interfere with the data of adjacent lines and cause incorrect or difficult delimitation on the peer device. The long strings of consecutive 0s or 1s must be processed to appear randomly generated before signals are carried over a line. This is the purpose of scrambling.

Scrambling generally involves inserting a fixed-length sequence at the local end and removing the sequence at the remote end. This inserted sequence keeps the signals stochastic over a line.

Trellis Coding

Common path coding techniques can be classified into convolutional coding and block coding. Trellis coding is a code modulation technique that combines convolutional coding with the digital modulation mode. The corresponding decoding technique is called Viterbi decoding.

The process of Trellis coding entails the redundancy of only one bit. Hence, Trellis coding features a higher coding efficiency and a simplified coding mechanism. However, the corresponding Viterbi decoding has a complicated process. Viterbi decoding can be divided into hard decision (HD) and soft decision (SD). SD adds some probability weighted calculation to the decoding process and thus Viterbi decoding has a stronger error correcting capability.

Trellis coding is mainly targeted at burst errors. It can correctly parse the discrete error bits in the transmission and features strong code gaining and error correcting capabilities. The VDSL2 standard defines Trellis coding as mandatory for VDSL2 implementation.

FEC

In general, there are multiple error correction mechanisms. Some depend on the transmission system itself to check the data and correct the errors after the data arrives at the peer end. Others only check the data and do not correct the errors; if any error is detected, the data is retransmitted. Forward error correction (FEC) belongs to the former category and applies to real-time services, as such services do not tolerate the latency caused by retransmission. FEC is not exclusive to DSL and is commonly used for error correction.

When applied in DSL, FEC uses Reed-Solomon (RS) coding and appends redundancy bytes to the original data. These redundancy bytes identify and correct errors. All error correction mechanisms have a trade-off in performance; accordingly, FEC sacrifices some bandwidth when implemented.

6 VDSL2 Access

About This Chapter

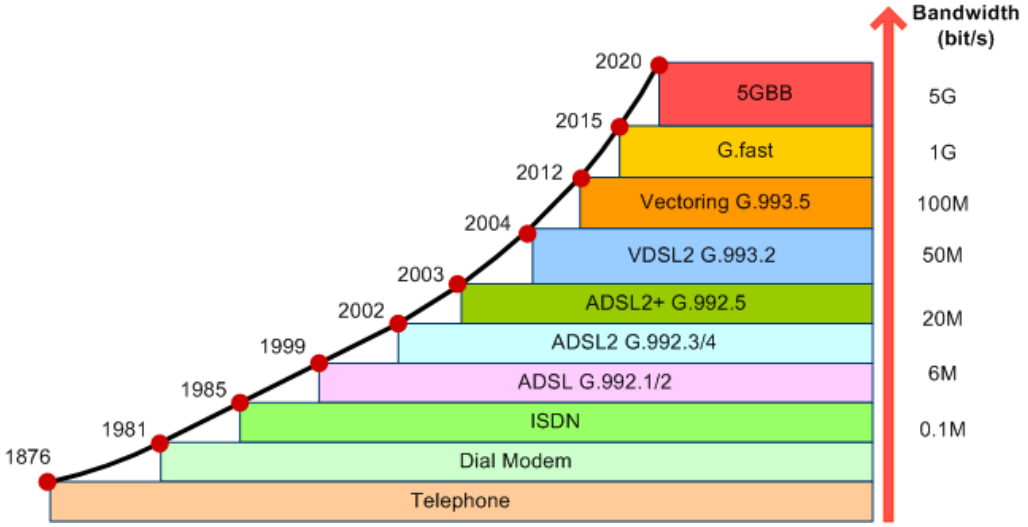
Very-high-speed digital subscriber line 2 (VDSL2) provides symmetric or asymmetric high-speed access services over twisted pairs. It increases the upstream/downstream access rates to symmetric 100 Mbit/s over a short distance (within 300 m), addressing the requirements for bandwidth-critical services such as high-definition (HD) video services. VDSL2 enables digital subscriber line access multiplexers (DSLAMs) to implement the "last mile" access, especially when the DSLAMs are deployed in fiber to the building/fiber to the curb (FTTB/FTTC) networks.

6.1 Overview of Mainstream Copper Line Technologies

VDSL2 and G.fast are mainstream copper line technologies. What are the positions of them in the copper line technology family? What are the highlights of VDSL2 and G.fast compared with other mainstream copper line technologies? Find the answers to these questions in this section.

The future broadband requirement is continuously increasing. Continuous technological innovation on copper lines (as shown in Figure 6-1) enables copper lines to meet the requirement of ultra-broadband network construction.

Figure 6-1 Copper line technology development



In network deployment, copper line access technologies include ADSL2+, VDSL2 (supporting vectoring to cancel inter-line crosstalk), and G.fast (supporting vectoring to cancel inter-line crosstalk), as shown in Table 6-1.

Table 6-1 Mainstream copper line technologies

Technology	Description	Parameter
	ADSL is a technology for transmitting high-speed private line services over common twisted pairs in asymmetric mode. ADSL2+ is an extension of ADSL and supports a maximum downstream rate of 24 Mbit/s, a maximum upstream rate of 2.5 Mbit/s, and a maximum transmission distance of 6.5 km.	• Typical rate: 128 kbit/s to 24 Mbit/s • Typical reach: longer than 1 km • Typical usage scenario: DSLAM/FTTC
	VDSL2 is an extension of VDSL1. VDSL2 is compatible with ADSL, ADSL2, and ADSL2+, but is not compatible with the less-common VDSL1.	• Typical rate: 30 Mbit/s to 50 Mbit/s (Rates can be improved to 50 Mbit/s to 100 Mbit/s after vectoring is enabled) • Typical reach: shorter than 1 km • Typical usage scenario: FTTB/FTTC **NOTE** Vectoring is a technology that uses vectoring algorithms to cancel crosstalk for multi-pair VDSL2 lines, thereby improving VDSL2 and G.fast line bandwidths.

Technology	Description	Parameter
	G.fast is a new high-bandwidth access technology applies to copper lines. It uses wider spectra than those used by ADSL/ADSL2+ and VDSL2 based on existing last-mile copper lines, helping carriers to rapidly deploy ultra-broadband networks by reusing existing infrastructure.	• Typical rate: 500 Mbit/s to 1 Gbit/s (with vectoring enabled) • Typical reach: shorter than 250 m • Typical usage scenario: FTTB/FTTD

6.2 VDSL2 Access Introduction

VDSL2 is based on ITU-T Recommendation G.993.2 and is an extension to VDSL1, which is based on ITU-T Recommendation G.993.1. VDSL2 is designed to be compatible with ADSL, ADSL2, and ADSL2+, but not the less-common VDSL1. VDSL2 features the following highlights:

The following are three prime drivers for VDSL2:

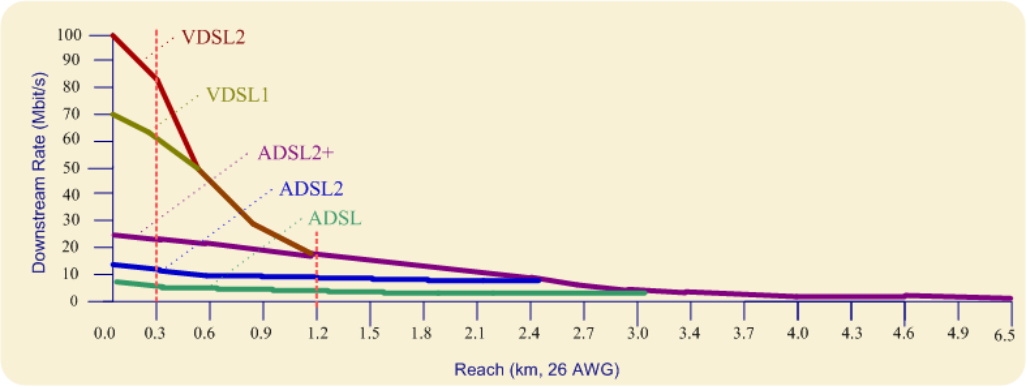
- Emergence of new broadband services: New broadband services, such as HDTV, require a higher access rate.
- Broadband network evolution facts: Copper-based access networks cannot evolve to full-fledged optical networks within a short time.
- Improvements in digital subscriber line (DSL) technology: DSL technologies have been advancing towards higher access quality, better user satisfaction, normalization among the DSL standards, and lower operating expenditure (OPEX).

VDSL2 features the following highlights:

- Higher access rate over short distances: VDSL2 stretches the spectrum range to 30 MHz and provides a symmetric 100 Mbit/s for upstream/downstream within 300 m, addressing the requirements for bandwidth-intensive services such as HDTV. VDSL2 typically applies to the "last mile" access of DSLAMs, especially for FTTB/FTTC access solutions.
- Higher transmission rate over longer distances: Compared with VDSL1, VDSL2 extends the spectrum and improves the transmit power spectrum density (PSD) to provide a higher transmission rate over longer distances.
- Compatibility with ADSL, ADSL2, and ADSL2+ terminals: VDSL2 supports packet transfer mode (PTM) 64/65-byte encapsulation based on IEEE 802.3ah, and asynchronous transfer mode (ATM) encapsulation used by ADSL, ADSL2, and ADSL2+. Therefore, VDSL2 is compatible with ADSL, ADSL2, and ADSL2+ terminals.
- Enhanced operation and maintenance (O&M) capabilities: VDSL2 supports line diagnosis and the acquisition of essential line parameters by dedicated line test procedures.

Figure 6-2 shows a comparison between VDSL2 and ADSL/ADSL2/ADSL2+/VDSL1 in terms of downstream rate and reach. Note that some DSL performance parameters, such as line activation rate, are associated with the electrical attributes of twisted pairs. Specifically, the smaller core diameter of a twisted pair means larger line attenuation. The following figure uses the common 26AWG twisted pair (core diameter: 0.4 mm) as an example.

Figure 6-2 Comparison between VDSL2 and ADSL/ADSL2/ADSL2+/VDSL1 in terms of downstream rate and reach



The preceding figure shows that:

- VDSL2 provides a remarkably higher downstream rate than ADSL/ADSL2/ADSL2+/VDSL1 within a 0.3 km reach. VDSL2, however, provides the theoretical 100 Mbit/s downstream rate only when the reach is within 0.25 km.
- VDSL2 provides the same downstream rate as VDSL1 and ADSL2+ at a 1.2 km reach.
- VDSL2 produces the same rate curve as ADSL2+ at a reach over 1.2 km.

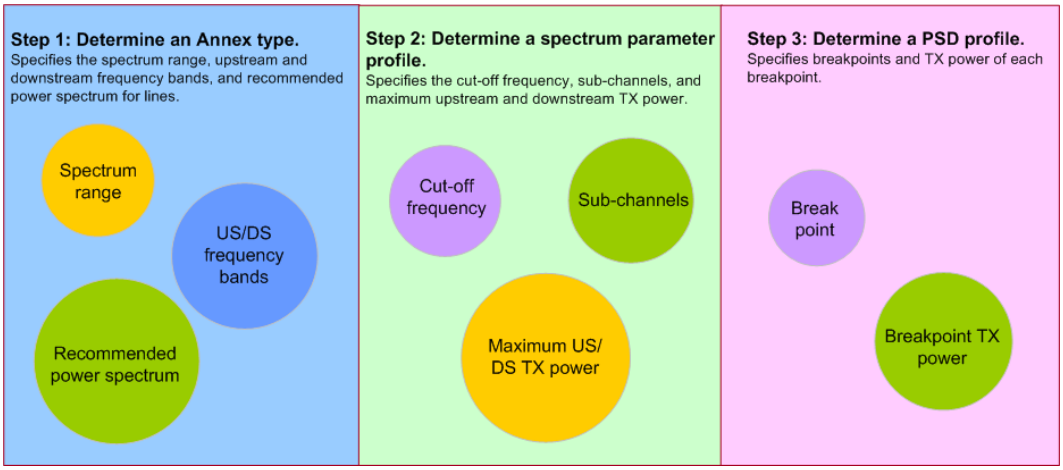
6.3 Basic VDSL2 Technologies

6.3.1 Overview of VDSL2 Spectrum Planning

The factors affecting DSL loops may vary depending on network conditions, and it is difficult to address the application requirements of different scenarios using a single mechanism. To account for this, the spectrum plan is split into two parts: the upstream/downstream band and power spectrum plan (based on *Annex type* and *PSD profile*, respectively), and the spectrum parameter plan (based on the *spectrum parameter profile*). A flexible combination of the two plans produces different spectrum profiles to meet diverse application requirements. Select a proper Annex type, spectrum parameter profile, and PSD profile to configure a spectrum profile.

Figure 6-3 shows overall VDSL2 spectrum planning.

Figure 6-3 Overall VDSL2 spectrum planning



NOTE

Knowledge about the G.992.3, G.992.5, G.993.2, G.997.1, and TR165 standards helps you better understand the spectrum plan described in this section.

6.3.2 Annex Types and US/DS Frequency Band Planning

Most DSL standards provide a generic definition in the body, and then a description about specific schemes in the Annex. The schemes specify how to use the low frequency band in typical application scenarios. The schemes also specify how to plan the upstream/downstream band (apart from the low frequency band) for data transmission and how to plan the power spectrum.

Users can select a proper Annex type by running commands. When an Annex type is selected, the upstream/downstream band plan and power spectrum plan are determined.

NOTE

The power spectrum plan is critical for controlling the performance and reliability of DSL lines. VDSL2 provides flexible power spectrum control mechanisms. The concepts and features related to the power spectrum plan are described in 6.3.6 PSD Profiles. As an Annex type includes a power spectrum plan, this section will also include information about power spectrum. It is recommended that you also read 6.3.6 PSD Profiles to better understand the VDSL2 feature.

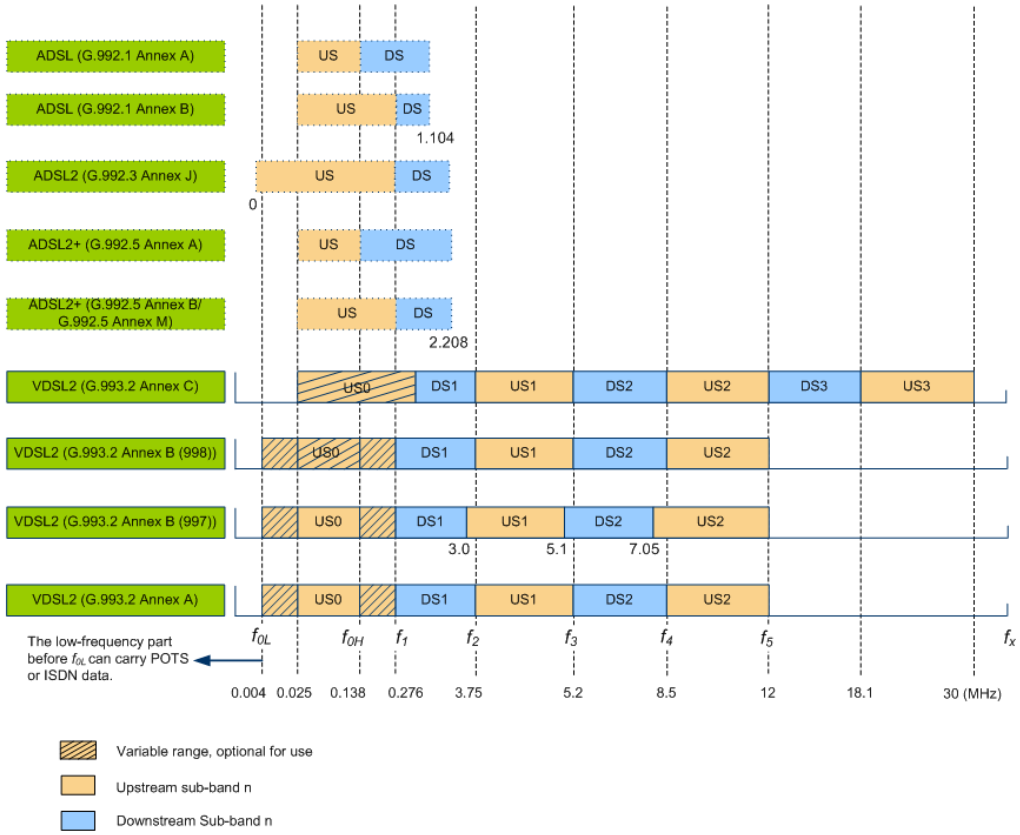
Annex Types and Upstream/Downstream Band Plans

An Annex type defines the scheme for using the low frequency band (the frequency band before f_{0L} as shown in Figure 6-4, used for carrying POTS or ISDN data) and the scheme for planning the upstream/downstream band (apart from the low frequency band) for data transmission. The upstream/downstream band plan specifies the spectral segments for upstream/downstream transmission, and the start and stop frequencies in each segment.

The spectral segment used for upstream transmission is called upstream sub-band (US), such as US0 and US1 in Figure 6-4; the spectral segment used for downstream transmission is called downstream sub-band (DS), such as DS1 and DS2 in Figure 6-4. The total number of USs and DSs in the entire band is the total number of bands specified in the spectrum profile. For example, "5 Band" indicates that the entire band is divided into five sub-bands.

For ADSL/ADSL2/ADSL2+, the entire available spectrum is divided into one US and one DS, as shown in Figure 6-4. This figure also shows mapping between US0 for VDSL2 and US for ADSL2+. The mapping is also described in 6.3.7 Limit PSD Mask.

Figure 6-4 ADSL2+/VDSL2 upstream/downstream band plan



Among the upstream sub-bands, US0 is optional (as shown in Figure 6-4) and an Annex type defines the frequency range of US0 (start frequency f_{0L}; stop frequency f_{0H}) and usage of US0. A Huawei access device also provides commands for enabling and disabling US0 and specifying a PSD mask.

For long-distance access, the upstream high frequency band is fully exploited, so the low frequency band becomes a valuable resource. Enabling US0 in this case will extend the DSL coverage and improve upstream line performance. VDSL2 can be activated beyond 1.4 km only when US0 is enabled. Usually, you are recommended to enable US0 beyond 800 m.

6.3.3 Command Parameters for US/DS Frequency Bands

This section describes Annex types and command parameters planned for upstream and downstream frequency bands.

Basic Parameters

Different DSL standards define different numbers of Annex types, some of which may even be empty. Annex types sharing the same name may contain different contents. For example, Annex A defined in ITU-T Recommendation G.992.5 differs from Annex A defined in ITU-T Recommendation G.993.2. An Annex type not designated with the standard number is meaningless.

When configuring spectrum profiles using commands, you can specify only a standard (that is, the standard used to establish a DSL link between the access device and its interconnected

modem); in this case, all the Annex types included in the standard are selected. Or, you can specify a standard and select some Annex types under this standard using the **Custom** parameter. When the latter method is used, the selectable Annex types and the standard are displayed on the CLI. The selected standard and Annex types determine the **Transmission Mode** for the DSL line.

- 1-T1.413
- 2-G.992.1 (Annex A/B/C)
- 3-G.992.2 (Annex A/C)
- 4-G.992.3 (Annex A/B/I/J/L/M)
- 5-G.992.4 (Annex A/I)
- 6-G.992.5 (Annex A/B/I/J/M)
- 7-G.993.2 (Annex A/B/C)
- 8-ETSI

NOTE

VDSL2 line parameters can be used in different combinations based on profiles. The configuration modes can be classified as TR129 (also called the common mode), TI, and TR165. For a Huawei access device, the default configuration mode is TR129. Carriers can switch between the configuration modes by running the **switch vdsl mode to** command. Considering the current development trend, it is recommended that you use TR165, which is more flexible than the others. The command parameters included in the following VDSL2-related topics are specific to the TR165 mode.

Table 6-2 lists the common xDSL standards and Annex types defined in each standard.

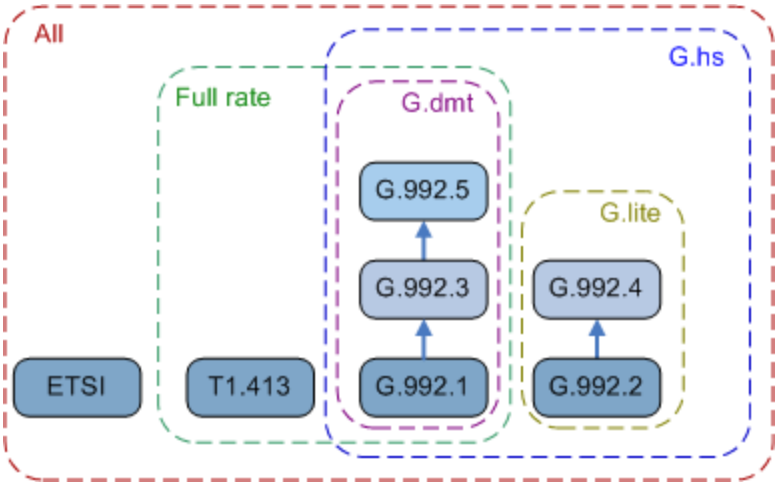
Table 6-2 Standards and Annex types

Category	Standard	Annex Type	Remarks
ADSL series standards	G.992.1	Annex A Annex B Annex C	The following describes the Annex types of ADSL series standards. • Annex A is also called ADSL over POTS; the low frequency band carries voice services. • Annex B is also called ADSL over ISDN; the low frequency band carries ISDN services. • Annex C is not supported by the Huawei access device. • Annexes I and J are "all-digital" mode. Only data services are carried but low-frequency services are not. Annex I has the same band plan as Annex A, and Annex J has the same band plan as Annex B. Annex I applies when the adjacent pair of a DSL line carries POTS services; Annex J applies when the adjacent pair of a DSL line carries ISDN services. Annex I is not supported by the Huawei access device. • Annex L is also called the reach extended ADSL2 (READSL2). Annex L uses fewer upstream/downstream bands but has a higher transmit power than Annex A. Higher transmit power helps extend the reach but also increases interference between lines. This characteristic restricts the use of Annex L.
	G.992.2	Annex A Annex C	
	G.992.3	Annex A Annex B Annex I Annex J Annex L Annex M	
	G.992.4	Annex A Annex I	
	G.992.5	Annex A Annex B Annex I Annex J Annex M	

Category	Standard	Annex Type	Remarks
			• Annex M extends the upstream band of Annex A and applies when high upstream rate is required. Practically, Annex types are selected based on the use of the low-frequency band planned for a DSL network. For example, voice services are widely used in North America and China, so Annex A is selected. This attribute makes Annex types region-specific.
	T1.413	-	-
	ETSI	-	-
VDSL2 standards	G.993.2	Annex A Annex B Annex C	1. The Annex types differ from those with the same names defined in the ADSL series standards. 2. G.993.2 Annex A specifies the band plan for North America, G.993.2 Annex B for Europe, and G.993.2 Annex C for Japan. This is why G.993.2 Annexes are also called "band plan for region". However, these Annex types are not restricted only to the listed regions; they differ mainly in that they define different upstream/downstream bands and power spectra.

On the CLI interface, the above-listed ADSL series standards are classified to ease configuration, as shown in the following figure.

Figure 6-5 Classification of ADSL series standards



In the figure above, "G.dmt" refers to the standards using discrete multi-tone (DMT) modulation technology; "G.lite" refers to the standards using half of the available spectrum; "Full rate" refers to the standard using the entire available spectrum; "G.hs" refers to the standards using G.994.1 for handshaking; "All" refers to all standards. According to this categorization, G.993.2 belongs to G.dmt, Full rate, G.hs, and All in command configuration.

You can select multiple standards and Annex types during configuration. The access device and its interconnected modem will negotiate to determine the optimal transmission mode for activating the line.

Advanced Parameters

After Annex types are specified, the Huawei access device configures a default frequency band planning mode, displayed by parameter **defmode** in the following terminal display, for each Annex type. Parameter **defmode** indicates all, including all frequency band planning modes. This parameter can only be modified.

In addition to parameter **defmode**, you can add a desired frequency band planning mode. To add a frequency band planning mode, do as follows:

1. Select **Y** when the system displays the message "Will you set mode-specific parameters?"
2. Press **1** and select the frequency band planning parameters to be added in the terminal display.

Optional frequency band planning modes comply with G.997.1 and support the parameters in the following terminal display.

```
>  Will you set mode-specific parameters? (y/n) [n]:y
>    Current configured modes:
>    1-defmode
>    Please select 1-Add  2-Modify  3-Save and quit   [3]:1

>    2-ansit1413                    3-etsi
>    4-g9921PotsNonOverlapped         5-g9921PotsOverlapped
>    6-g9921IsdnNonOverlapped         7-g9921IsdnOverlapped
>    8-g9921tcmIsdnNonOverlapped       9-g9921tcmIsdnOverlapped
>    10-g9922PotsNonOverlapped        11-g9922PotsOverlapped
>    12-g9922tcmIsdnNonOverlapped      13-g9922tcmIsdnOverlapped
>    14-g9921tcmIsdnSymmetric         15-g9923PotsNonOverlapped
>    16-g9923PotsOverlapped           17-g9923IsdnNonOverlapped
>    18-g9923IsdnOverlapped           19-g9924PotsNonOverlapped
>    20-g9924PotsOverlapped           21-g9923AnnexIAllDigNonOverlapped
>    22-g9923AnnexIAllDigOverlapped    23-g9923AnnexJAllDigNonOverlapped
>    24-g9923AnnexJAllDigOverlapped    25-g9924AnnexIAllDigNonOverlapped
>    26-g9924AnnexIAllDigOverlapped    27-g9923AnnexLMode1NonOverlapped
>    28-g9923AnnexLMode2NonOverlapped  29-g9923AnnexLMode3Overlapped
>    30-g9923AnnexLMode4Overlapped     31-g9923AnnexMPotsNonOverlapped
>    32-g9923AnnexMPotsOverlapped      33-g9925PotsNonOverlapped
>    34-g9925PotsOverlapped           35-g9925IsdnNonOverlapped
>    36-g9925IsdnOverlapped           37-g9925AnnexIAllDigNonOverlapped
>    38-g9925AnnexIAllDigOverlapped    39-g9925AnnexJAllDigNonOverlapped
>    40-g9925AnnexJAllDigOverlapped    41-g9925AnnexMPotsNonOverlapped
>    42-g9925AnnexMPotsOverlapped      43-g9932AnnexAPots
>    44-g9932AnnexAIsdn               45-g9932AnnexBPots
>    46-g9932AnnexBIsdn               47-g9932AnnexCPots
>    48-g9932AnnexCIsdn
>    Please select  [2]:
```

NOTE

- The preceding terminal display is only an example. Use the terminal display on the CLI of the Huawei access device.

- Dozens of parameters are involved because an Annex type may define multiple frequency band planning modes. For example, G.993.2 Annex B defines two frequency band planning modes, Plan 997 and Plan 998, as shown in Figure 6-4. After G.993.2 is amended, Annex B supports the following frequency band planning modes more: 997E17, 997E30, 998E17, 998E30, 998ADE17, 998ADE30, HPE17, and HPE30.

6.3.4 Annex Types and Power Spectrum Planning

The upstream/downstream band plan is closely related to the power spectrum plan, which is critical to performance control and reliability assurance for DSL lines.

Each Annex in the ADSL series standards and VDSL2 standard defines the upstream/downstream band plan and provides suggestions on the power spectrum plan.

Power spectrum plans are referred to as PSD profiles. For details on related concepts and features, see 6.3.6 PSD Profiles.

6.3.5 Spectrum Parameter Profiles

ITU-T Recommendation G.993.2 defines eight spectrum parameter profiles: 8a, 8b, 8c, 8d, 12a, 12b, 17a, and 30a, which specify different spectrum parameters. Spectrum parameter profiles are used with Annex types defined in ITU-T Recommendation G.993.2. Spectrum parameter values vary with the Annex types.

Definition

Spectrum parameter profiles are exclusive to VDSL2 and they are defined in ITU-T Recommendation G.993.2. ADSL series standards do not support spectrum parameter profiles. Spectrum parameter profiles are referred to as "profiles" in ITU-T G.993.2, and as "G.993.2 profiles" or "VDSL2 profiles" on the access device.

Table 6-3 lists the key parameters in the eight spectrum parameter profiles specific to "Annex B (998E)". For detailed meanings of each parameter, see "Profiles" in ITU-T Recommendation G.993.2.

Table 6-3 Key parameters in spectrum parameter profiles

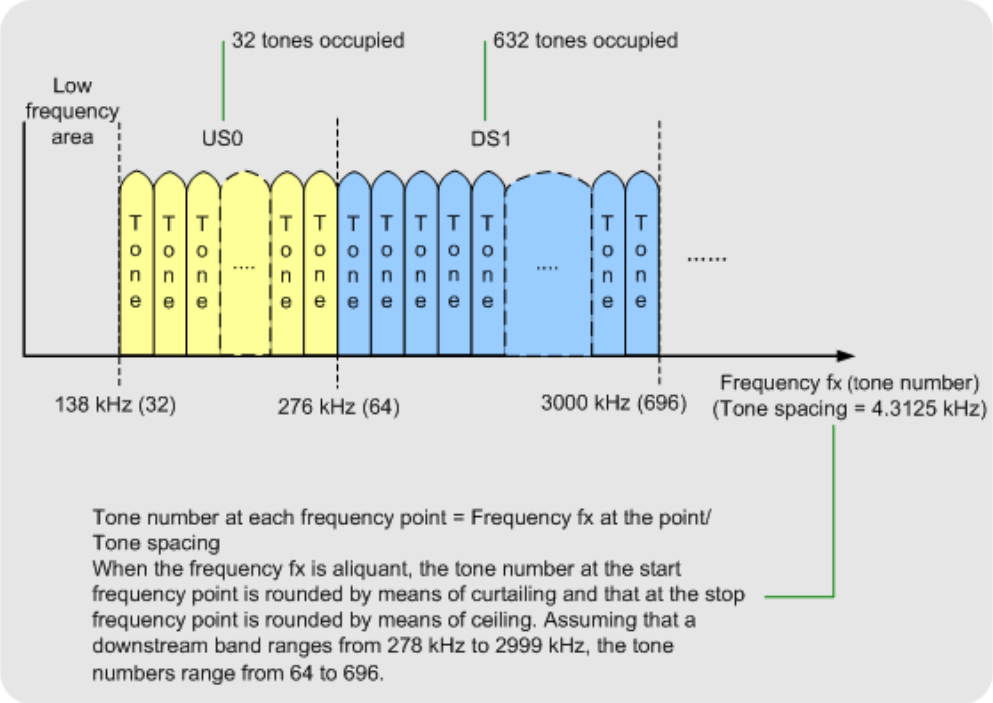
Profile	8a	8b	8c	8d	12a	12b	17a	30a
Bandwidth (MHz)	8.5	8.5	8.5	8.5	12	12	17.664	30
Tones	1972	1972	1972	1972	2783	2783	4096	3479
Tone spacing (kHz)	4.3125	4.3125	4.3125	4.3125	4.3125	4.3125	4.3125	8.625
Maximum aggregate downstream transmit power (dBm)	+17.5	+20.5	+11.5	+14.5	+14.5	+14.5	+14.5	+14.5
Maximum aggregate upstream	+14.5	+14.5	+14.5	+14.5	+14.5	+14.5	+14.5	+14.5

Profile	8a	8b	8c	8d	12a	12b	17a	30a
transmit power (dBm)								
Support of upstream band zero (US0)	Required	Required	Required	Required	Required	Regional annex dependent	Regional annex dependent	Not Supported

The following describes meanings of each parameter in Table 6-3.

- "Bandwidth" indicates the maximum stop frequency in the power spectrum used by the profile. The numbers in profile names indicate the parameter values of "bandwidth". For example, 12a and 12b indicate the maximum stop frequency of 12 MHz.
- The letters in profile names distinguish the "maximum aggregate downstream transmit power" attribute. For example, 8b indicates the maximum aggregate downstream transmit power of +20.5 dBm and 8c indicates +11.5 dBm. The maximum aggregate upstream transmit power of the eight profiles is the same (+14.5 dBm).
- As shown in Figure 6-6, VDSL2 uses the discrete multi-tone (DMT) technology, which divides the entire spectrum band into n tones (also called sub-carriers). In Table 6-3, "tones" indicates the number of tones in the entire spectrum band and "tone spacing" indicates the width of each tone.

Figure 6-6 VDSL2 tone division (for a band with f_{0L} of 138 kHz)



- **Support of upstream band zero (US0)** indicates whether the profile applies to the US0 band. Specifically, **Required** means that the profile applies to the US0 band, **Regional annex dependent** means that the profile may apply to the US0 band, depending on the region, and **Not Supported** means that the profile does not apply to the US0 band.

Applications

- The 8a and 8b profiles define high downstream transmit power and they typically apply to long-distance (800 m to 1000 m) VDSL2 application. The 8b profile defines a downstream transmit power of 20.5 dBm, which is the same as that defined for ADSL2+.
- Among the eight profiles, the 8c profile defines the lowest downstream transmit power (11.5 dBm) and it typically applies to VDSL2 in remote-end outdoor cabinets (distance range < 300 m; high access rate not required).
- The 8d and 12a/12b profiles define medium downstream transmit power and they typically apply to medium-distance (300 m to 800 m) VDSL2 application.
- The 17a and 30a profiles define a high stop frequency and, because of the high line attenuation, they typically apply to short-distance (< 300 m; high access rate required) VDSL2 application. The use of the 30a profile is restricted. This is because the 30a profile achieves the expected rate only in lab environment or when the line is short (< 150 m) and in good conditions. The 17a profile is hence more widely used.

This section provides only suggestions on applications of VDSL2 profiles and the user must select an appropriate profile depending on network conditions. Table 6-4 lists typical configurations for some commonly used profiles (with 26AWG twisted pairs of a 0.4 mm core diameter).

Table 6-4 Typical configurations for some commonly used profiles (with 26AWG twisted pairs of a 0.4 mm core diameter)

VDSL2 Profile	Activation Distance	Maximum Upstream Activation Rate	Maximum Downstream Activation Rate	US0 Enabled or Not
17a	<300 m	15 Mbit/s	50 Mbit/s	Yes
12a	300 m to 500 m	10 Mbit/s	40 Mbit/s	Yes
8b and 12a	500 m to 800 m	5 Mbit/s	25 Mbit/s	Yes
8b	800 m to 1000 m	2 Mbit/s	20 Mbit/s	Yes

6.3.6 PSD Profiles

The power spectrum plan is a PSD profile that defines the PSD masks for upstream or downstream frequency bands.

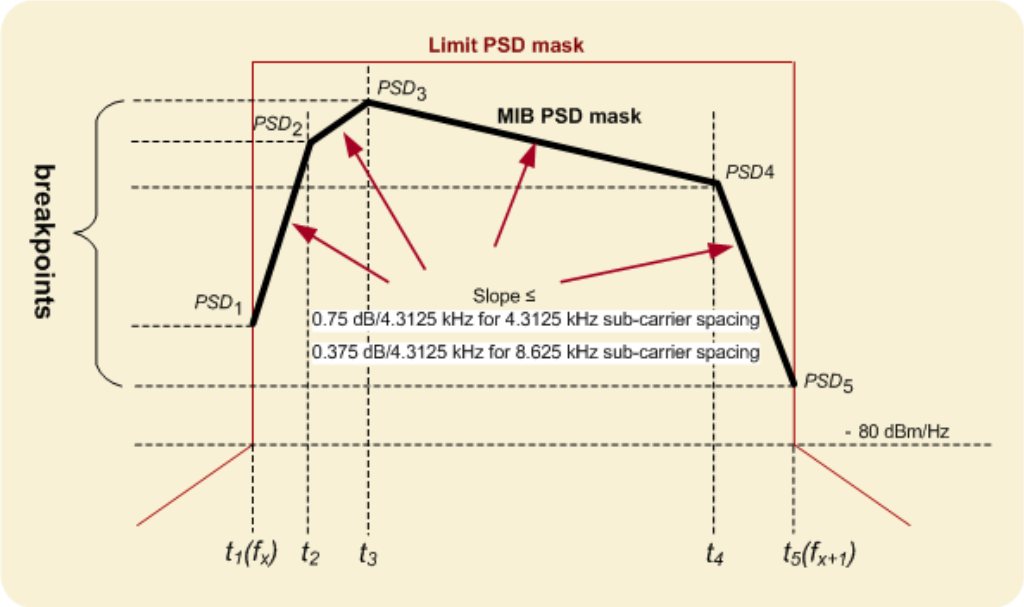
In this document, PSD profiles include not only TR165-specified mode-specific PSD profiles but also PSD contents specified in line spectrum profiles. Mode-specific PSD profiles are specified using "Please select the **mode specific PSD profile** index" of the **xdsl line-spectrum-profile add** command output.

The following two concepts are important for PSD profiles:

1. PSD refers to the differential coefficient of the transmit power at the frequency point and reflects the power intensity (expressed in dBm or Hz) at each frequency point. Users can derive the transmit power used in a spectrum band by performing integral calculation for PSD at each frequency point in the band. Controlling the PSD of a VDSL2 line protects the line against external noise and reduces the interference output of the line.
2. PSD mask is a fold line that links the **maximum PSD** at each frequency point. The system specifies PSD values for a series of breakpoints on a spectrum band and outlines the PSD mask of the spectrum band through an interpolation algorithm.

ITU-T Recommendation G.993.2 defines two types of PSD masks: limit PSD mask and management information base (MIB) PSD mask. Figure 6-7 shows the relationship between the two types of PSD masks. Both PSD masks can be configured by commands. The smaller PSD mask at each frequency point prevails.

Figure 6-7 Relationship between limit PSD mask and MIB PSD mask



NOTE

In Figure 6-7, the form of the limit PSD mask indicates that the MIB PSD mask should always be below the limit PSD mask (if the MIB PSD mask is above the limit PSD mask, the system chooses the smaller one as the PSD mask). The turns at the PSD mask in actual application cannot form a right angle, because the slope for each turn is restricted.

6.3.7 Limit PSD Mask

The limit power spectrum density (PSD) mask is defined in each Annex and is named **LIMITMASK** in the standard.

"Limit" indicates the highest PSD mask in a specified upstream/downstream band plan. The management information base (MIB) PSD mask must be lower than the limit PSD mask.

Limit PSD masks in upstream/downstream band plans vary with the Annex types. Table 6-5, Table 6-6, and Table 6-7 list the selectable limit PSD masks for Plan 998 (and its extensions) defined in ITU-T Recommendation G.993.2 Annex B.

NOTE

Since ITU-T Recommendation G.993.2 is continuously updated, Table 6-5, Table 6-6, and Table 6-7 may not be up to date and they are intended only as an explanation of the basic concepts of limit PSD masks.

Table 6-5 European limit PSD mask options for band plan 998 (and its extensions)

Short Name	Limit PSD Mask (Long Name)	Frequency	
		US0 Type A/B/M (see NOTE)	Highest Used Upstream or Downstream Frequency (kHz)
B8-1	998-M1x-A	A	12000
B8-2	998-M1x-B	B	12000
B8-3	998-M1x-NUS0	N/A	12000
B8-4	998-M2x-A	A	12000
B8-5	998-M2x-M	M	12000
...	...	...	...
B8-8	998E17-M2x-NUS0	N/A	17664
...	...	...	...
B8-11	998ADE17-M2x-A	A	17664
...	...	...	...
B8-13	998E30-M2x-NUS0	N/A	30000
...	...	...	...

NOTE – The US0 types stand for:

- US0 type A corresponds to Annex A/G.992.5;
- US0 type B corresponds to Annex B/G.992.5;
- US0 type M corresponds to Annex M/G.992.3/G.992.5;
- US0 type N/A designates a band plan variant that does not use US0.

The following describes the meanings of each field in Table 6-5.

- In Table 6-5, "Short Name" refers to the shortened name of a limit PSD mask and is used as an index. B8 indicates Plan 998. Similarly, B7 in Annex B indicates Plan 997. Carriers usually use short names to specify PSD masks. Table 6-6 and Table 6-7 define the breakpoints in each limit PSD mask and the PSD value at each breakpoint in the upstream and downstream directions, respectively. In the upstream direction (or the VTU-R transmit direction), the limit PSD mask of each upstream band applies; in the downstream direction (or the VTU-O transmit direction), the limit PSD mask of each downstream band applies. The following tables provide simplified contents regarding VTU-R limit PSD masks for band Plan 998. To view the complete contents, see ITU-T Recommendation G.993.2. Note that the upstream and downstream limit PSD masks defined in G.993.2 Annex B have the same short names, though they have different

breakpoints. The upstream and downstream limit PSD masks defined in G.993.2 Annex A have different short names (for upstream, examples include EU-32 and ADLU-32, as listed in "US0 type" below; for downstream, examples include D-32 and D-64).

Table 6-6 VTU-R limit PSD masks for band Plan 998 (and its extensions)

Name	**B8-1**	**...**	**B8-8**	**...**	**B8-11**	**...**
Long name	**998-M1x-A**	**...**	**998E17-M2x-NUS0**	**...**	**998ADE17-M2x-A**	**...**
kHz	**dBm/Hz**	**...**	**dBm/Hz**	**...**	**dBm/Hz**	**...**
0	-97.5	...	-100	...	-97.5	...
4	-97.5	...	-100	...	-97.5	...
4	-92.5	...	-100	...	-92.5	...
25.875	-34.5	...	-100	...	-34.5	...
50	-34.5	...	-100	...	-34.5	...
80	-34.5	...	-100	...	-34.5	...
120	-34.5	...	-100	...	-34.5	...
138	-34.5	...	-100	...	-34.5	...
...	...	...	...	...	...	...
24890	-100	...	-100	...	-100	...
25065	-100	...	-100	...	-100	...
30000	-100	...	-100	...	-100	...
30000	-110	...	-110	...	-110	...
30175	-110	...	-110	...	-110	...
≥30175	-110	...	-110	...	-110	...

Table 6-7 VTU-O limit PSD masks for band Plan 998 (and its extensions)

Name	**B8-1**	**...**	**B8-8**	**...**	**B8-11**	**...**
Long name	**998-M1x-A**	**...**	**998E17-M2x-NUS0**	**...**	**998ADE17-M2x-A**	**...**
kHz	**dBm/Hz**	**...**	**dBm/Hz**	**...**	**dBm/Hz**	**...**
0	-97.5	...	-97.5	...	-97.5	...
4	-97.5	...	-97.5	...	-97.5	...
4	-92.5	...	-92.5	...	-92.5	...
80	-72.5	...	-72.5	...	-72.5	...

Name	B8-1	...	B8-8	...	B8-11	...
Long name	998-M1x-A	...	998E17-M2x-NUS0	...	998ADE17-M2x-A	...
kHz	dBm/Hz	...	dBm/Hz	...	dBm/Hz	...
...	...	...	...	...	...	...
24890	-100	...	-100	...	-100	...
25065	-100	...	-100	...	-100	...
30000	-100	...	-100	...	-100	...
30000	-110	...	-110	...	-110	...
30175	-110	...	-110	...	-110	...
≥30175	-110	...	-110	...	-110	...

- In Table 6-5, "Long Name" describes a limit PSD mask. For example, 998 and 998E17 indicate band plans, which are mentioned in "Annex Types"; NUS0 indicates that US0 is disabled. Long names do not clearly show the specific PSD mask plan, which must be determined according to the detailed PSD mask definitions, such as Table 6-6 and Table 6-7, in the standard.
- In Table 6-5, "US0 Type" defines the US0 types associated with each limit PSD mask, as indicated by the note in Table 6-5.
  - Type A indicates that US0 has the same spectrum range as G.992.5 Annex A, that is, 25 kHz to 138 kHz.
  - Type B indicates that US0 has the same spectrum range as G.992.5 Annex B, that is, 120 kHz to 276 kHz.
  - Type M indicates that US0 has the same spectrum range as G992.3/G.992.5 Annex M, that is, 25 kHz to 276 kHz.
  - Type N/A indicates that US0 is disabled.
- In Table 6-5, the last column "Highest Used Upstream or Downstream Frequency" outlines the stop frequency in the spectrum associated with the limit PSD mask. The limit PSD mask defines PSD values for a series of breakpoints within the stop frequency range.

Table 6-5 shows a limit PSD mask plan for VDSL2. The ADSL series standards also define a spectrum plan, which includes one upstream sub-band and one downstream sub-band, and therefore the power spectrum plan is simple. G.992.3/ G.992.5 Annex M defines stop frequencies for 10 upstream bands (EU-32 to EU-128) and associated limit PSD mask profiles (upstream only). G.992.3/G.992.5 Annex J defines stop frequencies for ten upstream bands (ADLU-32 to ADLU-128) and associated limit PSD masks profiles (also upstream only). The numbers in "EU-32" and "ADLU-32" indicate the serial numbers of tones associated with the stop frequencies of the upstream band. For example, "32" indicates that the stop frequency of the upstream band maps 32 tones. Assuming that the tone spacing is 4.3125 kHz, then the stop frequency of the upstream band is 138 kHz. "EU" refers to the extended upstream sub-band and complies with the Annex M features; "ADLU" refers to all digital mode upstream sub-band and complies with the Annex J features.

According to the definition of "US0 type" in Table 6-5, the VDSL2 US0 band maps the ADSL US band. Therefore, the power spectrum profiles defined in the ADSL series standards also apply to VDSL2 US0. You can select these power spectrum profiles as needed when configuring the VDSL2 US0 PSD mask.

6.3.8 Command Parameters for Limit PSD Masks

This section describes command parameters limit PSD masks.

According to TR165, configuring a VDSL2 PSD profile involves configuring the US0 PSD mask and PSD masks for other bands.

1. To configure a US0 PSD mask, run the **xdsl line-spectrum-profile add** command, and then select option **y** at the prompt message "Will you set US0 PSD masks."
2. To configure PSD masks for other bands, run the **xdsl mode-specific-psd-profile add** command, and then select option **y** at the prompt message "Will you set VDSL2 limit PSD masks."

Many limit PSD masks (referred to as LIMITMASKs, in line with the standard) are optional. In order to simplify configurations, LIMITMASKs in G.997.1 are classified into PSD mask classes (or referred to as CLASSMASKs) according to the Annex types defined in G.993.2. Similarly, G.997.1 classifies spectrum parameter profiles into the following classes by stop frequency:

- Class 8: Profiles 8a, 8b, 8c, 8d
- Class 12: Profiles 12a, 12b
- Class 17: Profile 17a
- Class 30: Profile 30a

On the CLI interface, select the CLASSMASK for the desired LIMITMASK and specify LIMITMASK for each profile class. Table 6-8 lists mappings between CLASSMASK, profile class, and LIMITMASK.

Table 6-8 is organized as follows:

1. Lines 1, 2, and 3 show the CLASSMASK options for each Annex type.
2. For command configurations, the rarely-applied CLASSMASKs are grouped. For details, see the value range for "VDSL2 PSD mask class selection" in **xdsl mode-specific-psd-profile add**. The options for "LIMITMASK for each CLASSMARK" are shown in each row of Table 6-8. The following uses the first line in profile class 8 as an example to show the mappings between command parameter settings and contents in the table below (see the «-marked fields).

NOTE

The following command output is only an example. During actual configuration, the actual command output prevails.

```
>   VDSL2 PSD mask class selection:
>   1-Class 998 Annex A or Class 997-M1c Annex B or Class 998-B Annex C
>   2-Class 997-M1x Annex B or Class 998-CO Annex C
>   3-Class 997-M2x Annex B
>   4-Class 998-M1x Annex B
>   5-Class 998-M2x Annex B
>   6-Class 998ADE-M2x Annex B
>   7-Class HPE-M1 Annex B
>   Please select (1~7) [5]:1 «
>   Current LIMITMASK for each CLASSMASK you can choose:
```

```
>   Profile8a/b/c/d:
>   1: Limit1: D-32 M1c-A-7 POTS-138b«  2: Limit2: D-48 TCM-ISDN
>   3: Limit3: POTS_276b           4: Limit9: D-64
>   5: Limit10: D-128
>   Profile12a/12b:
>   6: Limit1: D-32 POTS-138b       7: Limit2: D-48 TCM-ISDN
>   8: Limit3: POTS_276b           9: Limit9: D-64
>   10: Limit10: D-128
>   Profile17a:
>   11: Limit1: D-32 POTS-138b      12: Limit2: D-48 TCM-ISDN
>   13: Limit3: POTS_276b          14: Limit9: D-64
>   15: Limit10: D-128
>   Profile30a:
>   16: Limit1: D-32 POTS-138b      17: Limit2: D-48 TCM-ISDN
>   18: Limit3: POTS_276b          19: Limit9: D-64
>   20: Limit10: D-128
>  Please select (1~20) [1]:
```

Table 6-8 Definition of LIMITMASK for each CLASSMASK

Profile Class	PSD Mask Classes									
	Annex A	Annex B							Annex C	
	998 Annex A «	998-M1x Annex B	998-M2x Annex B	998ADE-M2x Annex B	997-M1x Annex B	997-M1c Annex B «	997-M2x Annex B	HPE-M1 Annex B	998-B Annex C «	998-CO Annex C
8 «	D-32 «	M1x-A	M2x-A	M2x-A		M1c-A-7 «	M2x-A		POTS-138b «	POTS_138co
8	D-48	M1x-B	M2x-B	M2x-B	M1x-M-8		M2x-M-8		TCM-ISDN	POTS_276co
8			M2x-M	M2x-M	M1x-M		M2x-M		POTS_276b	
8		M1x-NUS0	M2x-NUS0	M2x-NUS0						
8	D-64									
8	D-128									
12	D-32	M1x-A	M2x-A	M2x-A			M2x-A		POTS-138b	POTS_138co
12	D-48	M1x-B	M2x-B	M2x-B					TCM-ISDN	POTS_276co
12			M2x-M	M2x-M	M1x-M		M2x-M		POTS	

Profile Class	PSD Mask Classes									
	Annex A	Annex B							Annex C	
	998 Annex A «	998-M1x Annex B	998-M2x Annex B	998ADE-M2x Annex B	997-M1x Annex B	997-M1c Annex B «	997-M2x Annex B	HPE-M1 Annex B	998-B Annex C «	998-CO Annex C
									_276b	
12		M1x-NUS0	M2x-NUS0	M2x-NUS0						
12	D-64									
12	D-128									
17	D-32		E17-M2x-NUS0	ADE17-M2x-A			E17-M2x-NUS0	17-M1-NUS0	POTS-138b	
17	D-48		E17-M2x-NUS0-M	ADE17-M2x-B					TCM-ISDN	
17				ADE17-M2x-NUS0-M					POTS_276b	
17	D-64									
17	D-128									
30	D-32		E30-M2x-NUS0	ADE30-M2x-NUS0-A			E30-M2x-NUS0	30-M1-NUS0	POTS-138b	
30	D-48		E30-M2x-NUS0-M	ADE30-M2x-NUS0-M					TCM-ISDN	
30									POTS_276b	
30	D-64									
30	D-128									

6.3.9 MIB PSD Mask

ITU-T Recommendation G.993.2 defines management information base (MIB)-controlled power spectrum density (PSD) masks for flexible control over PSD.

"MIB-controlled" means configuring PSD masks through the network management system (NMS) or through a digital subscriber line access multiplexer (DSLAM). MIB-controlled PSD masks provide users with more options than the limit PSD masks defined in standards. Carriers can control the power spectrum and reduce crosstalk by configuring suitable PSD masks according to DSLAM distribution, distance to users, and coexistence of ADSL and VDSL. Such user-configured PSD masks are referred to as MIB PSD masks.

For details on MIB PSD masks, see MIB-controlled PSD Mask.

6.4 Key VDSL2 Techniques

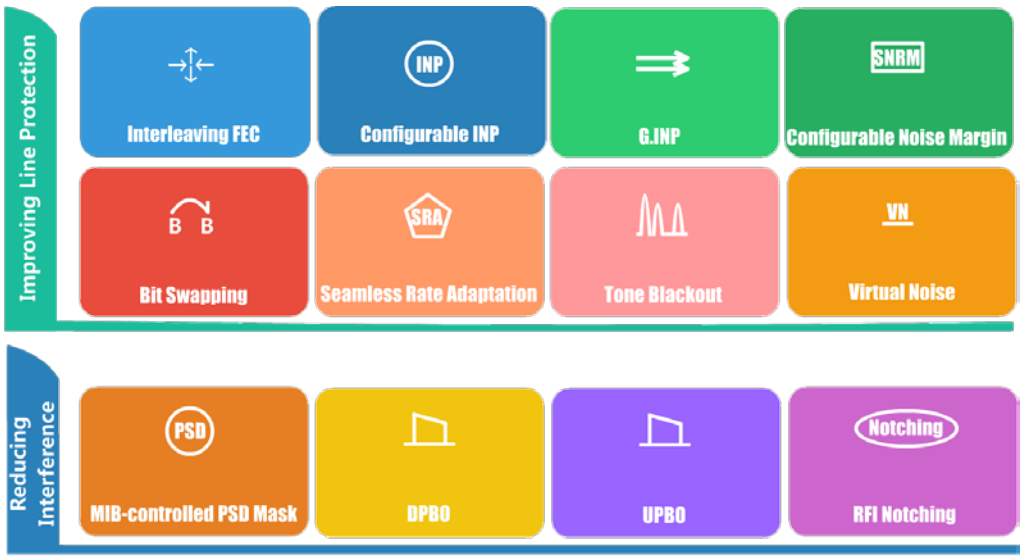
6.4.1 Overview of Key VDSL2 Techniques

This section provides an overview of key VDSL2 techniques for improving bandwidths and stability of VDSL2 lines.

Key VDSL2 techniques include:

- Techniques for improving line protection
- Techniques decreasing noise output
- VDSL2 PTM bonding

The preceding two types of techniques are briefly introduced as follows.



6.4.2 Key Techniques for Improving Line Protection

DSL provides various techniques for improving line protection, such as enhanced error detection and correction, reserved noise margin, and online reconfiguration (OLR). All the techniques employed translate into higher line stability.

Interleaving FEC

Forward error correction (FEC), though having powerful error correction capability, is insufficient for handling long strings of consecutive bit errors that are generated in severe line noise. Hence, interleaving FEC is introduced. Interleaving FEC is a major approach for avoiding pulse interference.

Working Principle of Interleaving FEC

Interleaving may be block interleaving or convolutional interleaving, and DSL uses the latter. Compared with convolutional interleaving, block interleaving is simple but less effective. The following uses block interleaving as an example to illustrate the interleaving process.

Figure 6-8 shows a typical interleaver. In this example, the rectangle block refers to an interleaving block and the numbers in the block indicate the sequence in which bits enter the interleaver. Generally, bits are written by row and read by column. The interleaving depth (D) is 3 and interleaving width (I) is 7. In practical applications, an interleaver has greater D and I values.

NOTE

ADSL directly uses the FEC codeword N_{FEC} as the interleaver width, whereas VDSL2 uses the fraction ($I = N_{FEC}/q$) of N_{FEC} as the interleaver width, with q ranging from 1 to 8.

Figure 6-8 Working principle of the interleaver

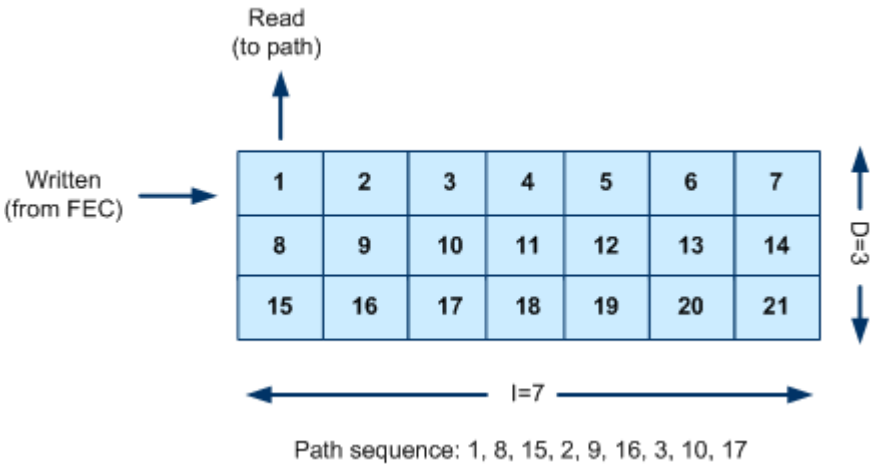


Figure 6-9 shows a de-interleaver that corresponds to the interleaver shown in Figure 6-8. The de-interleaver outputs cells in their correct sequence.

Figure 6-9 Working principle of the de-interleaver

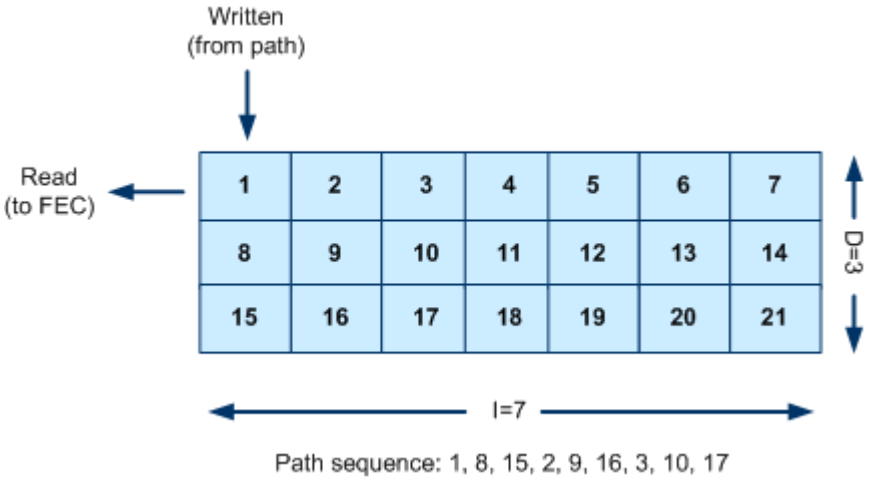


Figure 6-10 shows the benefit of interleaving by comparing the received bit errors with and without interleaving. In the figure, the first two rows indicate the sequence in which bits are transmitted over channels and the last two rows indicate the received bits. If a burst error similar to the third row occurs, bit errors will be distributed when interleaving takes effect so that they can be better corrected.

Figure 6-10 Comparison of received bit errors with and without interleaving

BITS sequence (without interleaving)	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
BITS sequence (with interleaving)	1	8	15	2	9	16	3	10	17	4	11	18	5	12	19	6
the burst errors																
the received BITS (without interleaving)	1	2				6	7	8	9	10	11	12	13	14	15	16
the received BITS (de-interleaving)	1		3	4	5	6	7	8		10	11	12	13	14		16

ITU-T Recommendation G.993.2 also defines a mechanism for dynamically adjusting the interleaving depth (D). In the handshake process, the office and user devices negotiate whether to support dynamic adjustment of the interleaving depth. If yes, the system adjusts the interleaving depth based on line conditions, thereby extending the range for SRA.

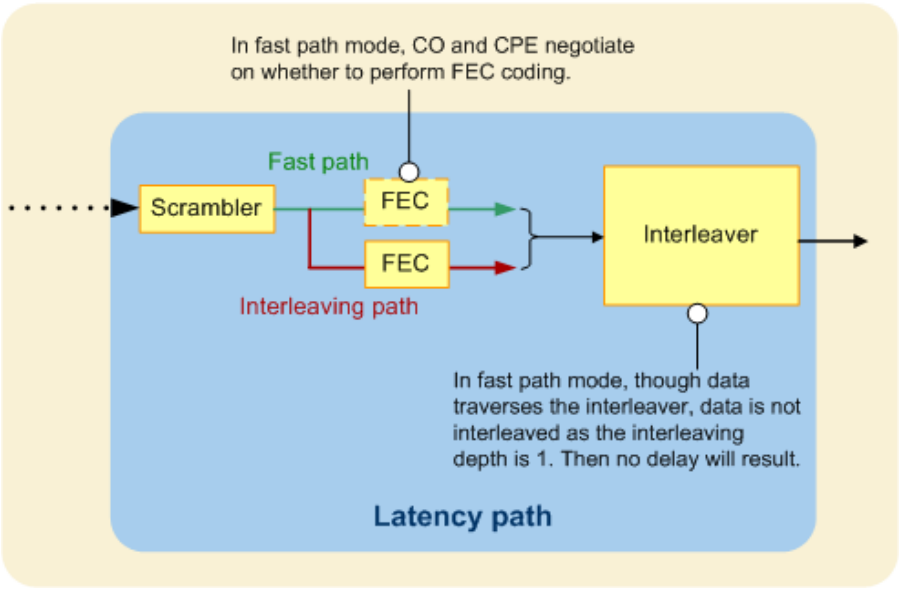
Path Mode and Maximum Interleaving Delay

Interleaving improves the line error correction capability by splitting consecutive bit errors on a line among various FEC frames. As the interleaving takes additional time, delay (referred to as interleaving delay) results. The **maximum interleaving delay** parameter is designed on a Huawei access device to control the interleaving delay. Specifically, the interleaving delay produced after a port is activated cannot exceed the maximum interleaving delay. On the Huawei access device, users can run the **xdsl inp-delay-profile add** command to set the maximum interleaving delay.

As interleaving delay will impact delay-critical services, such as VoD, voice, and fax services, VDSL2 allows users to select a path mode ("path" means "latency path" and has the same

meaning as the path in "dual-latency path") before line initialization: fast path or interleaving path. Figure 6-11 shows how the two path modes vary from each other.

Figure 6-11 Fast path and interleaving path



- Fast path: The line has a shorter delay but smaller error correction capability. In this mode, the interleaving depth is 1, which means no interleaving is performed, and the maximum interleaving delay is 0 ms.

NOTE

- ITU-T Recommendation G.997.1 defines three special values for the maximum interleaving delay:
- S0: **Interleaving delay** is set to **0**, indicating no limit on the maximum interleaving delay.
- S1: **Interleaving delay** is set to **1**, indicating the interleaving depth (D) of 1 and the maximum interleaving delay of 0 ms.
- S2: **Interleaving delay** is set to **255**, indicating the maximum interleaving delay of 1 ms.
- For the VDSL2 service boards in the H802 and H80A series, which agree with ITU-T G.997.1, set "interleaving delay" to 1 (S1 in ITU-T G.997.1) and INP to 0 to select the fast path mode; for the VDSL2 service boards in the H80B, H805, and H808 series, which use a different mechanism, set "interleaving delay" to 0 and INP also to 0 to select the fast path mode.

- Interleaving path: In interleaving path mode, the system has stronger error correction capability but a longer delay. It is typically applicable to the services that are not reliability or delay-critical, such as file download. In this mode, the FEC-processed bit stream is sent to the interleaver and then to the line. On the other side of the line, the bit stream is de-interleaved.

In VDSL2, the interleaving capability is represented by **interleaving depth (D)**, the error correction capability by **minimum INP** (see Configurable INP Parameters for details on INP), and interleaving delay by **maximum interleaving delay**, which are correlated to each other. In other words, deeper interleaving means more powerful error correction capability (greater INP value) but longer interleaving delay. The three parameters fit a formula defined in ITU-T Recommendation G.993.2.

In practical application, the system does not judge the minimum INP or maximum interleaving delay but applies the settings to a board directly. The board will make adaptation to ensure successful line activation after receiving the settings. Generally, use a longer interleaving delay (63 ms, for instance) if the minimum INP value is large (16, for instance).

If the minimum INP value is small and the maximum interleaving delay is short, the line will be activated with a low rate or probably cannot be activated.

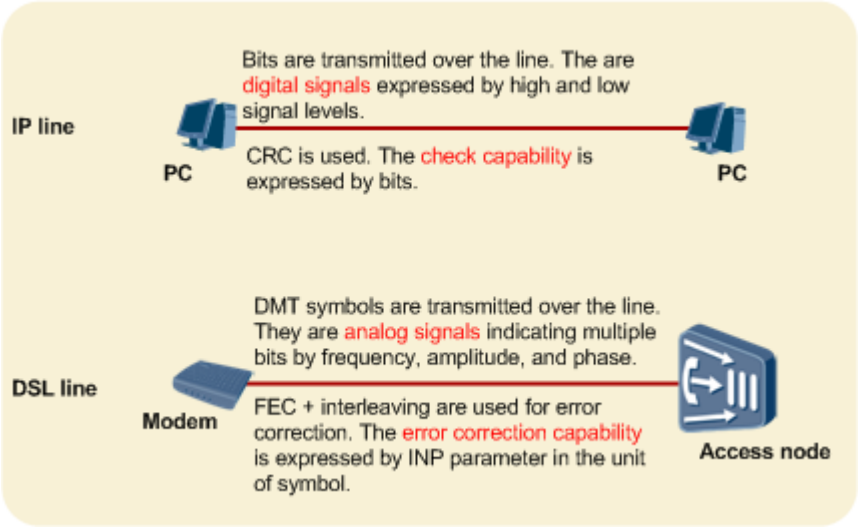
Configurable INP Parameters

Impulse noise protection (INP) refers to a technical category. In the DSL standard, INP indicates the error correction capability of a line or, more specifically, the count of correctable consecutive discrete multi-tone (DMT) symbols during de-interleaving.

INP Definition

Figure 6-12 shows the definition of INP parameters. On the device, **minimum INP** controls the error correction capability. The INP value of an activated port must be equal to or larger than **minimum INP**.

Figure 6-12 INP indication



The DMT symbol rate is an influence factor for conversion between INP parameter values and pulse noise duration. The DMT symbol rate is defined as 8000 DMT symbols per second in the 30a profile and as 4000 DMT symbols per second in other spectrum profiles. "INP=16" means that the system can correct the bit errors produced in the noise duration of 16 x 1/8000 = 2 ms in the 30a profile, and 16 x 1/4000 = 4 ms in other spectrum profiles.

INP Parameter Application

In ADSL2+/VDSL2, the interleaving capability is represented by **interleaving depth (D)**, the error correction capability by **minimum INP**, and interleaving delay by **maximum interleaving delay** (see Interleaving FEC for details on interleaving), which are correlated to each other. In other words, deeper interleaving means more powerful error correction capability (a greater INP value) but a longer interleaving delay. The three parameters fit a formula defined in ITU-T Recommendation G.993.2.

On the Huawei access device, users can run the **xdsl inp-delay-profile add** command to configure INP (or the interleaving delay). A board adjusts the interleaving depth and delay based on the specified minimum INP for the system to suppress pulse noise interference. If

erasure decoding is used, INP can be significantly increased without additional redundancy (no impact on the efficiency for carrying payload).

In practical application, the system does not judge the minimum INP or maximum interleaving delay before applying the settings to a board. The board will make adaptation to ensure successful line activation after receiving the settings. Generally, use a longer interleaving delay (63 ms, for instance) if the minimum INP value is large (16, for instance). If the minimum INP value is small and the maximum interleaving delay is short, the line will be activated with a low rate or probably cannot be activated. This means that there is a correlation between INP and the activated line rate. When the interleaving depth is constant, a greater INP value means a sharper decrease of the activated line rate.

When configuring the minimum INP, users must note the following conditions:

- If the Internet access rate is low, the line probably has a long delay. The most possible cause of the long delay is a large INP value.
- In the ADSL2+/VDSL2 over POTS service, there will be an abrupt change in line impedance after an onhook, producing transient pulse signals on the line. In this case, the ADSL2+/VDSL2 line will lose packets or even result in offline instances. It is recommended to set the minimum INP to 2 or greater for ADSL2+/VDSL2 over POTS.

The optimal INP value must be determined based on statistics of line noise distribution and spectrum range monitored over a long duration in order for the system to minimize the impact on line performance while maintaining a stable line. Impulse noise monitor (INM) is used for the monitoring.

Erasure Decoding

When used with FEC (Reed-Solomon coding), erasure decoding increases the system INP value without requiring additional redundancy.

Erasure decoding is optional as defined in the standard and the device manufacturers decide whether to implement it on central office (CO) and customer premises equipment (CPE) devices.

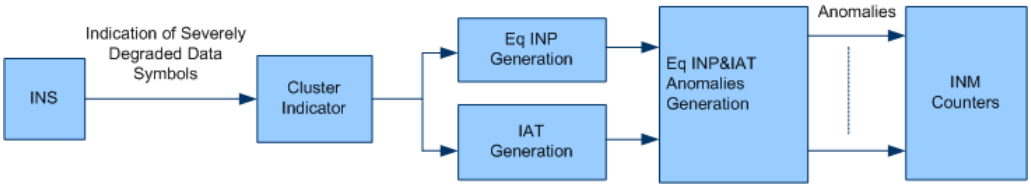
Impulse Noise Monitor (INM)

A greater INP value means more powerful line error correction capability, but longer data transmission delay and lower efficiency of carrying payload. Therefore, setting an optimal INP value is important to ADSL2+/VDSL2.

The optimal INP value must be determined based on statistics of line noise distribution and spectrum range monitored over a long duration in order for the system to minimize the impact on line performance while maintaining a stable line. Impulse noise monitor (INM) is used for the monitoring.

Figure 6-13 shows the working principle of INM.

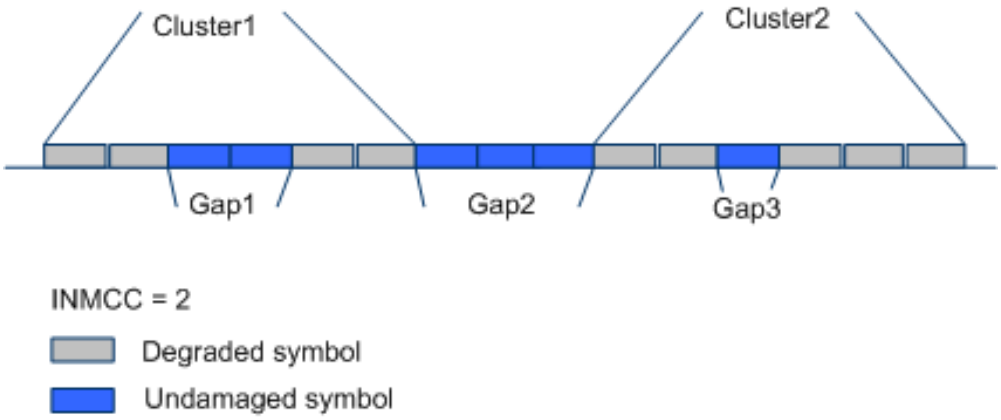
Figure 6-13 Working principle of INM



Working principle of INM:

1. The impulse noise sensor (INS) checks for severe damage in DMT symbols. If DMT symbols are severely damaged, they are downgraded.
2. The cluster indicator identifies INS-detected DMT symbols and groups the matched DMT symbols in a cluster. Clusters are preconditions for later DMT symbol processing. Figure 6-14 shows the process of identifying DMT symbols in clusters.

   Figure 6-14 Working principle of INM

   - As shown in the figure above, INM cluster continuation value (INMCC) is a key parameter for a cluster. INMCC indicates the maximum number of intact DMT symbols that can be included in a cluster. In this example, INMCC is 2 and Gap1 has two DMT symbols, which belong to a cluster (Cluster 1). Gap2 has three DMT symbols, higher than the limit. Therefore, Cluster1 includes only Gap1 and Gap2 does not belong to any cluster.
3. The Eq INP generation module calculates equivalent INP (INP_eq) for each cluster, and the inter arrive time (IAT) generation module calculates IAT for the entire symbol series. IAT refers to the number of symbols between two consecutive clusters, excluding the Sync symbol.
4. The Eq INP & IAT anomalies generation module collects statistics of **INP_eq** and **IAT**.
5. The INM counters count **INP_eq** and **IAT** by a certain rule, and produce irregular INP_eq and IAT bar charts based on the data. Users can view and use the data, and configure **INP_Min** (minimum INP) and **Delay_Max** (maximum interleaving delay) based on **INP_eq** and **IAT**.
6. Users can query the INM statistical results by running the **display statistics performance** command, or view the INP_eq and IAT bar charts using the NMS.

Physical Layer Retransmission (G.INP)

Some pulse noise may produce numerous bit errors. To protect a system against the pulse noise, one theoretical approach is to improve impulse noise protection (INP) by increasing forward error correction (FEC) redundancy and interleaving depth. However, the theoretical approach is not feasible because it causes a long delay and low efficiency in carrying payload, or has high requirements on components. ITU-T Recommendation G.998.4 defines physical layer retransmission to provide an alternative for improving INP. Specifically, physical layer retransmission improves INP while providing a high transmission rate and an acceptable transmission delay, and it is typically applicable to line quality-critical services, such as video services.

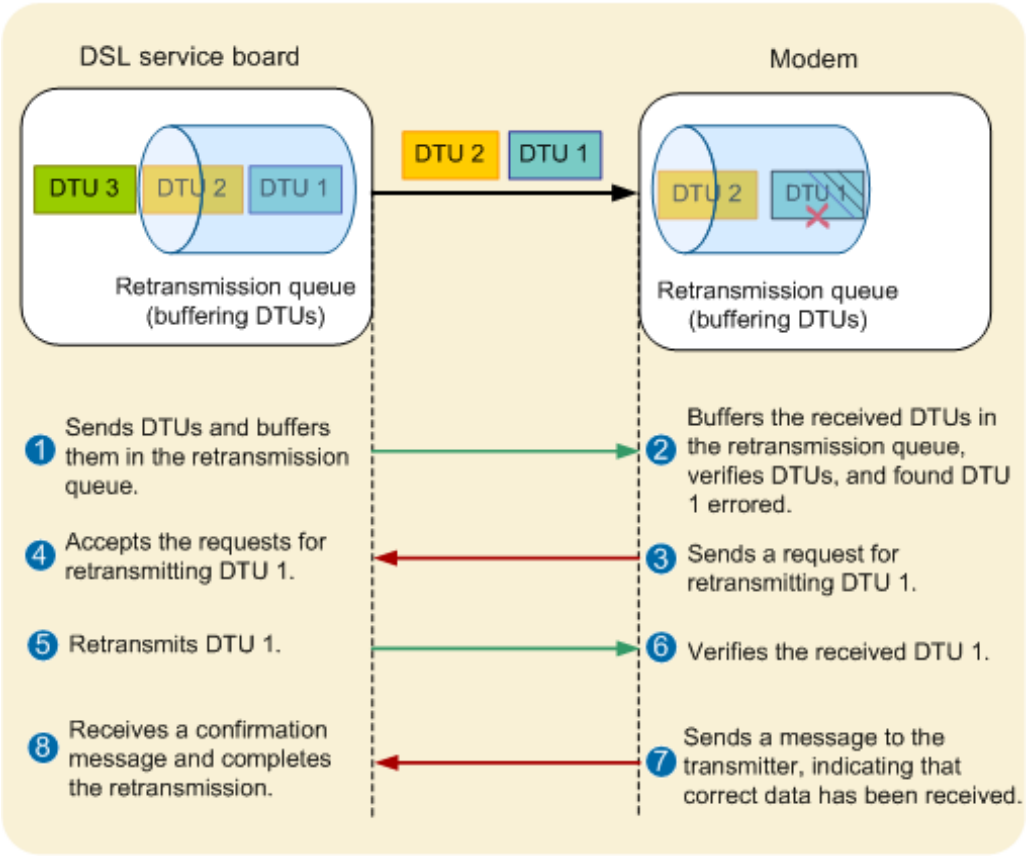
G.INP is another designation of ITU-T Recommendation G.998.4. Physical layer retransmission is referred to as RTX.

G.INP is intended to protect the system against the following types of pulse noise:

- Single high impulsive noise event (SHINE), which is neither repetitive nor periodic, but unpredictable because it is caused by burst impulse.
- Repetitive electric impulsive noise (REIN), which is repetitive and is caused by the electric main line and influenced by the local AC frequency.

Figure 6-15 shows how the access device implements retransmission in the downstream direction. Retransmission in the upstream direction is similar.

Figure 6-15 Working principle of retransmission



As shown in Figure 6-15, both the transmitter and receiver provide retransmission queues. To start the retransmission process, the transmitter encodes the to-be-sent data in data transfer units (DTUs), which are buffered in a retransmission queue. After receiving the DTUs, the receiver also buffers them in a retransmission queue and verifies them. If a DTU is found errored, the receiver sends a retransmission request to the transmitter. Then, the transmitter retransmits the DTU as requested. When receiving the retransmitted DTU, the receiver verifies it. If the DTU is correct, the receiver sends an acknowledgement message to the transmitter. By now, the retransmission process is completed.

In line with ITU-T Recommendation G.998.4, the Huawei access device supports G.INP retransmission parameter settings. For details, see G.998.4-related parameters in the **xdsl line-spectrum-profile add**, **xdsl inp-delay-profile add**, and **xdsl data-rate-profile add** commands. Users can query statistics of retransmission performance and operation specifications by running the **display xdsl statistics performance**, **display line operation**, and **display channel operation** commands.

Configurable Noise Margin

Noise margin is also signal-to-noise ratio (SNR) margin. The line conditions, such as ambient temperature, humidity, and ambient background noise, keep changing, and so does the SNR of each tone. A noise margin is retained when bits are allocated to each tone. When the line conditions change, the SNR decreases. If the SNR decrease is within the noise margin, the bit error ratio (BER) can stay lower than the standard-stipulated 10^{-7}, and data can be properly transmitted.

Concepts

Noise margin

Noise margin refers to the extra noise that the access device can tolerate while retaining the existing rate and BER. A wider noise margin means a more stable line but a lower activated physical connection rate.

Bit allocation

The noise power spectrum and line attenuation vary with the frequency, and different tones have varied SNRs and number of allocated bits. Therefore, different tones have varied noise margins but only one noise margin value is displayed. In practical application, the lowest noise margin will apply as the noise margin of the entire xDSL line.

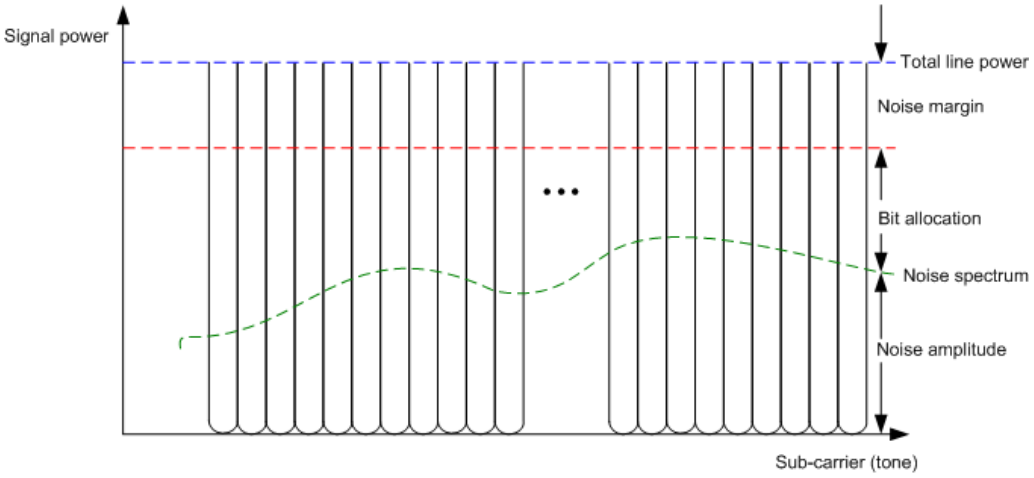
SNR

As a basic indicator in the communication industry, SNR reflects path quality. SNR refers to the ratio of the energy of data signals carried over each tone to the noise energy. Therefore, the xDSL SNR is the SNR of each tone. Each tone's signal and noise energy is expressed in dBm/Hz. Noise power ranges from -120 dBm/Hz to -140 dBm/Hz, and signal transmit power ranges from -40 dBm/Hz to -90 dBm/Hz. A tone with a 3 dB SNR can carry one bit. For a tone to carry 15 bits, the tone must have an SNR of at least 45 dB.

Working Principle

Figure 6-16 shows how noise margin works. Each tube represents a tone, the blue line represents total line power, the area outlined by the blue and red lines represents the reserved noise margin, and the area below the green line represents noise power. As shown in the figure, the area outlined between the red and green lines is used for carrying transmission signals (bit allocation).

Figure 6-16 Noise margin



When no noise margin is reserved, a noise amplitude increase may push the total signal power over the blue line, producing bit errors or even user offline events. When noise margin is reserved, the access device can tolerate a certain noise amplitude increase, allowing the total signal power to stay between the blue and red lines. In this way, the access device achieves higher line stability.

Application

The activated noise margin is associated with the target noise, and maximum and minimum noise margins configured for the access device. Specifically, the activated noise margin is close to the target noise margin, and within the range outlined by the maximum and minimum noise margins. A higher reserved noise margin means less power for carrying bits and a lower transmission rate.

Noise margins, including target, maximum, and minimum noise margins, apply in both upstream and downstream directions.

Target noise margin

- Target noise margin refers to the noise margin required for an access device to initialize with a BER of 10^{-7} or smaller. The target noise margin applies during line training and does not take effect after a line is trained. The line must be initialized with a BER of 10^{-7} or smaller. After line training is complete, users can query the actual noise margin of the line, which is close to the target noise margin.
- The target noise margin is reserved during normal data communication and it ensures normal communication in unfavorable line conditions. A larger noise margin means a less probability for the access device to encounter data transmission errors, a safer access device, but a lower maximum rate. For practical applications, configure a proper target noise margin based on line conditions.
- The access device establishes xDSL line connections and determines their rates according to the target noise margin. An over-high target noise margin may cause a decrease in the activated line rate, and an over-low target noise margin may cause an unstable line.

Maximum noise margin

- For a line in good conditions, if the activated noise margin exceeds the maximum noise margin, the access device must lower the line SNR by decreasing the signal power, while retaining the line rate up to the line requirement.
- In the process of xDSL connection establishment, if the noise margin calculated by the access device exceeds the specified maximum noise margin, the port will lower the signal power so that the noise margin will decrease to lower than the maximum noise margin.

Minimum noise margin

- When the line conditions turn unfavorable and the activated noise margin is lower than the minimum noise margin, the line cannot carry the expected bits. In this case, the line SNR must be raised by increasing the signal power so that the line can provide the required rate. If the signal power cannot be increased at all or cannot be increased to the extend to push the noise margin higher than the minimum noise margin, the line must be retrained.
- In the process of xDSL connection establishment, if the calculated noise margin is lower than the preset minimum noise margin, the port fails to be activated.

Determine the maximum and minimum noise margins based on line conditions. The maximum and minimum noise margin settings apply after the line is activated. A line keeps changing, sometimes in a good way and sometimes in a bad way.

- When the line condition worsens and the noise margin is lower than the minimum noise margin, the line cannot carry the expected bits. In this case, the line SNR must be raised by increasing the signal power so that the line can provide the required rate.
- When the line condition improves and the noise margin is higher than the maximum noise margin, the line SNR is over-high and will result in resource waste. In this case, the SNR must be lowered by decreasing the signal power, while the required line rate is retained.

An over-high target noise margin may decrease the activated rate, while an over-low target noise margin may result in an unstable line. Retain the default value (6 dB) for the target noise margin generally. If the activated rate is required at 0 km, the target noise margin can be reduced to a certain extent, but it is recommended that you retain the value greater than 3 dB; otherwise, the line may be unstable. In other conditions, the default value is recommended.

Bit Swapping

Bit swapping automatically adjusts the bit and power allocation on different tones according to SNR changes, so that the line is dynamically adaptive to variable noise without retrainings.

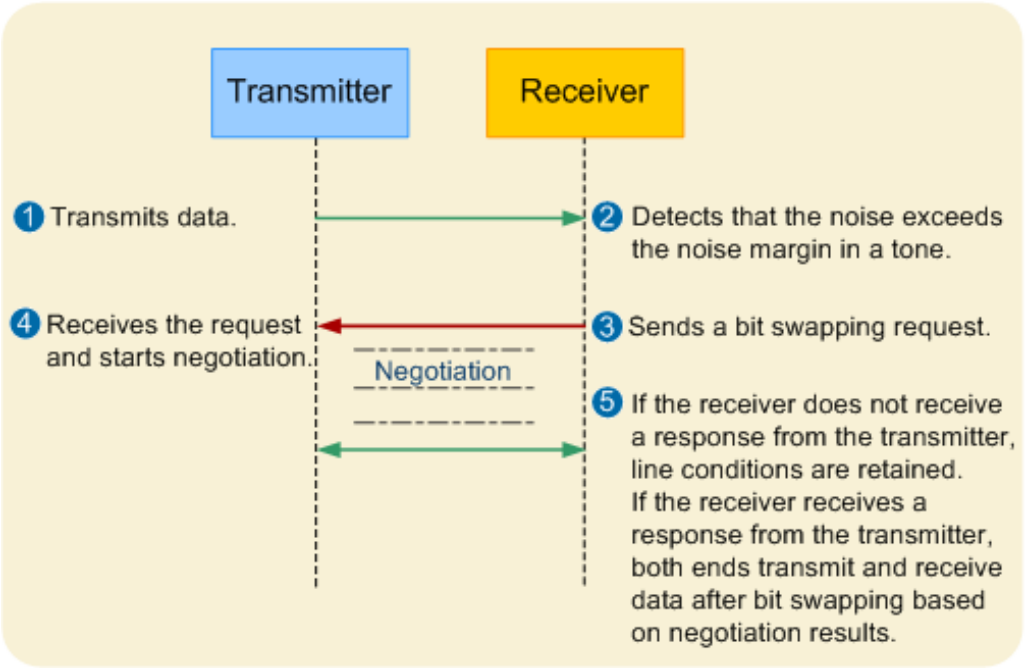
When the DSL line SNR changes but does not exceed the noise margin, the line BER meets the requirement (lower than 10^{-7}). However, noise margin does not always apply. When the line SNR decreases below the noise margin, the line BER will exceed 10^{-7}, and if it lasts for a long time, the line will be retrained to be adaptive to the noise.

Bit swapping automatically adjusts the bit and power allocation on different tones according to SNR changes, so that the line is dynamically adaptive to variable noise without retrainings.

As an online reconfiguration (OLR) technique, bit swapping does not change the line rate.

Figure 6-17 shows the working principle of bit swapping.

Figure 6-17 Working principle of bit swapping



- When detecting noise exceeding the noise margin in a tone, the receiver sends requests to the transmitter, requesting the transmitter to: swap bits from low-SNR tones to high-SNR tones; reduce the transmit power of the tones with reduced bits (crosstalk will result if these tones retain the original transmit power); increase the transmit power of the tones with increased bits.
- After the receiver sends bit swapping requests, the transmitter and receiver negotiate. Specifically, if the receiver does not receive response within a certain period of time, it deems that the transmitter does not support bit swapping (for example, when bit swapping is disabled) and retains the line conditions. If the receiver receives response from the transmitter, the transmitter and receiver will operate based on the negotiation results, to transmit or receive data. As devices (especially modems) supplied by different manufacturers have varied implementation of bit swapping, the transmitter and receiver, while negotiating and interacting with each other, may misunderstand each other. When misunderstanding happens, the line may be deactivated.

The Huawei access device allows users to enable or disable bit swapping in the upstream and downstream directions by running the **xdsl line-spectrum-profile add** command.

SRA

Bit swapping adjusts bit distribution on tones for a line to be noise-adaptive while retaining a constant rate. Seamless rate adaptation (SRA) enables the line to dynamically adapt to noises to a greater extent without retrainings.

When line conditions turn unfavorable and bit swapping fails to retain the bit error ratio (BER) at the required level, SRA decreases the rate; when line conditions turn favorable again, SRA increases the rate. In this manner, bandwidth usage is maximized.

Concepts

Association between line rates and bits

Line rate refers to the sum of bits transmitted over all tones on a channel.

SNR margin for rate upshift: When the noise margin reaches the specified value and sustains for **minimum upshift time**, the transmission rate automatically upshifts. **SNR margin for rate upshift** can be specified separately in upstream and downstream directions.

SNR margin for rate downshift: When the noise margin reaches the specified value and sustains for **minimum downshift time**, the transmission rate automatically downshifts. **SNR margin for rate downshift** can be specified separately in upstream and downstream directions.

Minimum upshift time: If the signal-to-noise ratio (SNR) margin reaches the value where the transmission rate starts to upshift, the transmission rate holds at this point for the specified minimum time and upshifts. **Minimum upshift time** can be specified separately in upstream and downstream directions.

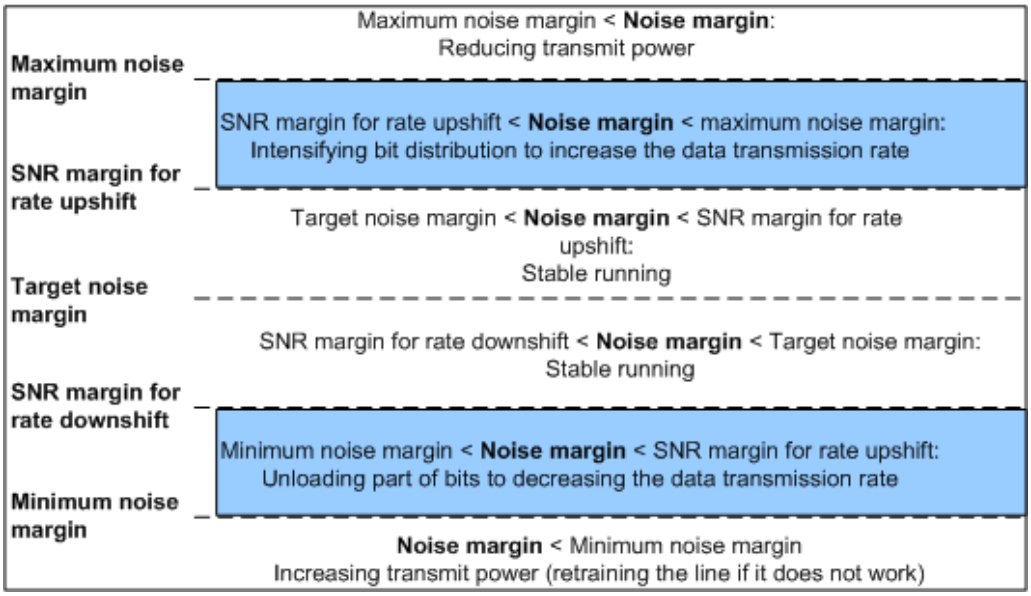
Minimum downshift time: If the SNR margin reaches the value where the transmission rate starts to downshift, the transmission rate holds at this point for the specified minimum time and downshifts. **Minimum downshift time** can be specified separately in upstream and downstream directions.

Working Principle

Figure 6-18 shows the association between a noise margin and SRA. The green-shaded blocks include description of SRA functions and the noise margin range.

- When **noise margin** is greater than or equal to **SNR margin for rate upshift** for over **minimum upshift time**, SRA functions to intensify bit distribution on the line for the transmission rate (line rate) to upshift.
- When **noise margin** is less than or equal to **SNR margin for rate downshift** for over **minimum downshift time**, SRA functions to unload part of bit distribution on the line for the transmission rate (line rate) to downshift.
- When **noise margin** is less than **SNR margin for rate upshift** but greater than **SNR margin for rate downshift**, or stays shorter than the minimum time, SRA will not function.

Figure 6-18 Noise margin

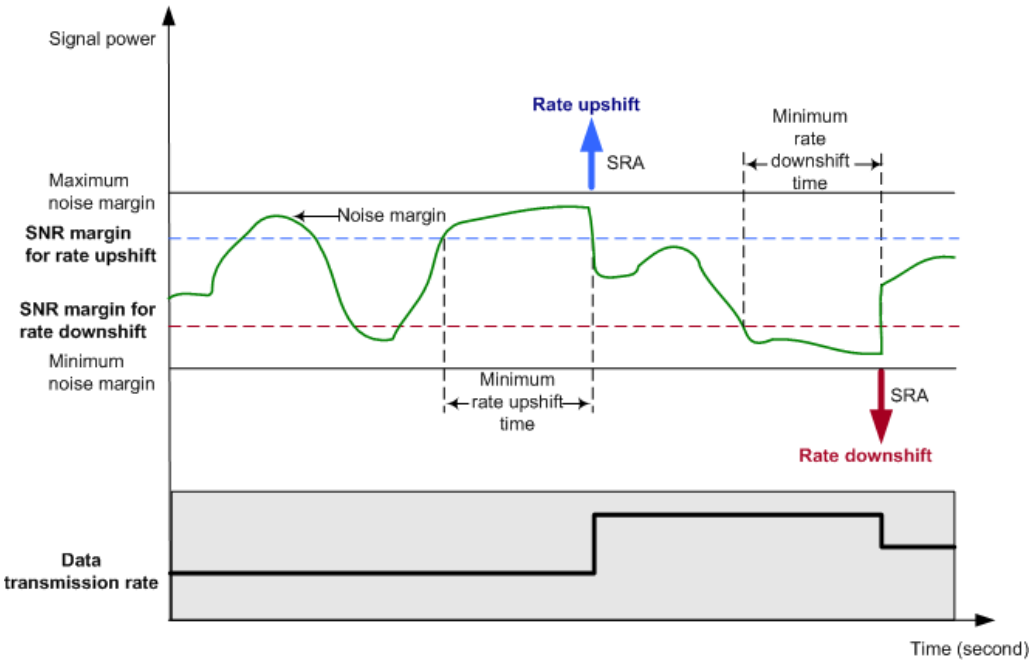


- When the noise margin is decreasing to lower than the SNR margin for rate downshift (which lies between the minimum and target noise margins), the customer premises equipment (CPE) sends control messages to the central office (CO), requesting the CO to dynamically decrease the signal transmit rate. After the signal transmit rate downshifts, the line noise margin increases. When the noise margin increases to the target value, the signal transmit rate stays stable.
- When the noise margin is increasing to higher than the SNR margin for rate upshift (which lies between the maximum and target noise margins), the CPE sends control messages to the CO, requesting the CO to dynamically increase the signal transmit rate. After the signal transmit rate upshifts, the line noise margin decreases. When the noise margin decreases to the target value, the signal transmit rate stays stable.

The rate upshift and downshift do not cause line retrainings or service interruption. This is why the rate adaptation process is seamless.

Figure 6-19 shows the entire SRA process and the specific process where the CO controls SRA using parameters.

Figure 6-19 SRA process



The rate does not upshift or downshift immediately when the line noise margin reaches the SNR margin for rate upshift or downshift. Instead, SRA starts to function only after the line noise margin stays at the level for the required time (in a range of 0s to 16383s).

Application

SRA can be enabled or disabled for an activated line. The receiver (CPE) triggers SRA while the transmitter (CO) controls SRA parameters.

SRA is sufficient to resolve the issues caused when noise margin changes slowly, but is insufficient when noise margin changes sharply.

SOS

Save our showtime (SOS) is a technology for enhancing line stability. Compared with seamless rate adaptation (SRA), SOS features faster line stability detection, which significantly reduces port offline rates caused by sudden noise increase. In addition, line gains remain unchanged during the entire SOS process, preventing unstable noise increase to lines.

When loud noises are suddenly increased to lines, the SOS feature allows the ports to work at a rate lower than before without going offline, which minimally affects services and supports rapid service recovery. After the noises are eliminated, the SOS feature allows the ports to work at a rate as before to recover the lines.

SOS Process Parameters

The SOS feature complies with the G.993.2 standard. Table 6-9 shows the parameters involved in an SOS process.

Table 6-9 Parameters involved in an SOS process

Parameter	Description	Corresponding Command Parameters
SOS-TIME	Indicates the SOS time window. If the value of this parameter is 0, SOS is disabled.	*sos-window-ds* *sos-window-us* For details, see the **xdsl sos-profile quickadd** command.
SOS-NTONES	Indicates the threshold for the percentage of degraded tones.	*sos-percent-degraded-tones-ds* *sos-percent-degraded-tones-us* For details, see the **xdsl sos-profile quickadd** command.
SOS-CRC	Indicates the threshold for the number of abnormal CRCs.	*sos-min-crc-ds* *sos-min-crc-us* For details, see the **xdsl sos-profile quickadd** command.
MAX-SOS	Indicates the maximum number of SOS processes.	*max-sos-ds* *max-sos-us* For details, see the **xdsl sos-profile quickadd** command.
MIN-SOS-DR	Indicates the minimum data rate of a valid SOS request.	**min-sos-dr** For details, see the **xdsl data-rate-profile quickadd** command.

SOS Rules

The SOS feature obeys the following rules:

1. The **SOS-TIME** value cannot be 0.
2. During the time specified by **SOS-TIME**, if the number of abnormal CRCs received by the receive end is greater than **SOS-CRC**, or the system determines that the percentage of degraded tones is greater than **SOS-NTONES**, the system triggers an SOS process.
3. If the number of SOS processes within 120s is greater than **MAX-SOS**, the modem switches to work in L3 state. If the line rate is continuously lower than **MIN-SOS-DR** for 20s, the modem also switches to work in L3 state.

SOS Process

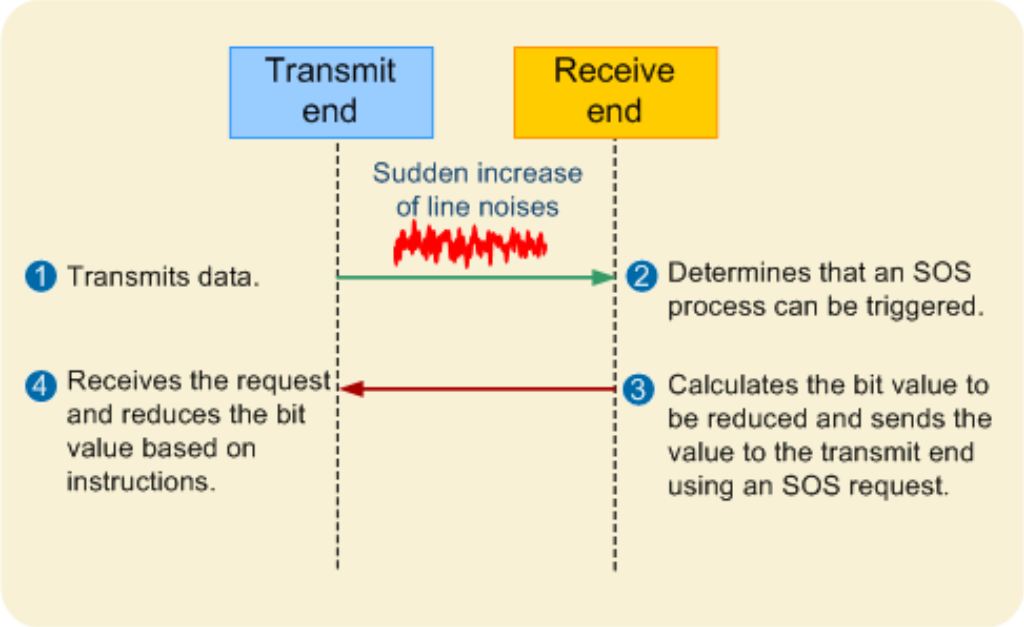
The SOS feature divides the subcarriers used by the VDSL2 system into multiple subcarrier groups. When line noises suddenly increase and an SOS process is triggered, the entire SOS process is as follows:

1. The receive end sends an SOS request that carries a simple and short message, notifying the transmit end of the bit value to be reduced. During the entire SOS process, the gain

remains unchanged. The SOS request is transmitted over a robust overhead channel (ROC), a logical channel dedicated for transmitting overhead messages.

2. Based on the received message, the transmit end reduces the bit value allocated to all subcarriers in the subcarrier group, preventing a large number of data exchanges between the transmit and receive ends for the bit and gain values allocated to each subcarrier.

Figure 6-20 SOS process



The entire SOS process is complete within several hundred milliseconds, at least one order of magnitude faster than the SRA process. In addition, the retained gain prevents the introduction of new unstable noises to lines.

Tone Blackout

If a certain band on the DSL line has unstable noise, which may cause interference, tone blackout can forbid the band from transmitting data, hence eliminating the interference. Some bands may be used for special purposes in certain regions; to prevent interference with these bands, tone blackout can forbid these bands.

Tone blackout, or missing tone as called in ADSL standards, means that a subcarrier is disabled and it will not carry any power (though there is a negligible transition band at both ends of the blackout band, because of the analog components), or any bit.

On the Huawei access device, users can run the **xdsl line-spectrum-profile add** command to configure tone blackout. The tone blackout band cannot be over-extensive or include the pilot tone; otherwise, the line may fail to be activated.

NOTE

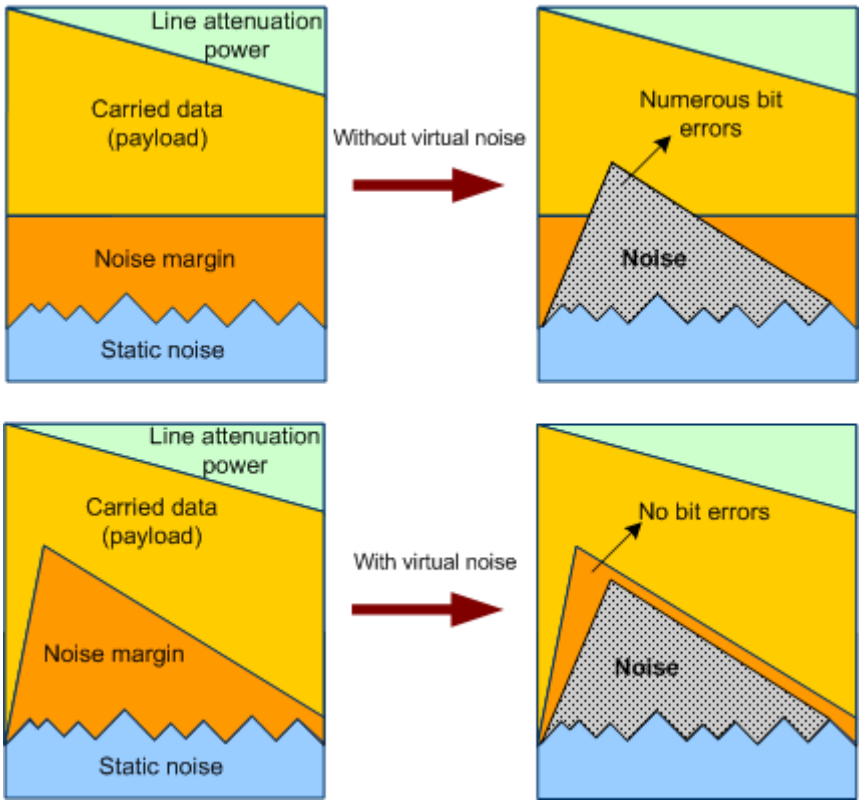
The system determines the pilot tones in line with ITU-T Recommendation G.994.1. Users can identify the pilot tones by comparing the spectrum profile against the ITU-T Recommendation G.994.1. Generally, the tone blackout band has a high frequency while the pilot tone has a low frequency, and they are less likely to intersect.

Virtual Noise

Noise margin is constant but line noise changes (the change fits a function of frequency). An over-large noise margin means fewer bits carried over tones and compromised performance; an over-small noise margin means a high BER when noise of a tone exceeds the noise margin. To resolve the issues, the noise margin power spectral density (PSD) mask must resemble the noise PSD mask whenever possible. This is how virtual noise helps.

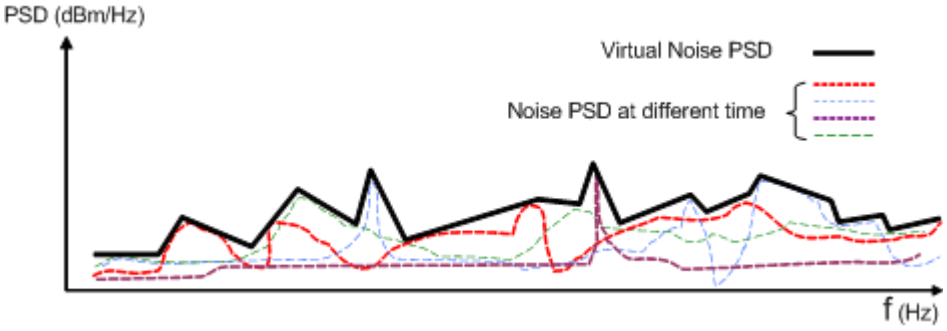
Figure 6-21 shows a reference model of virtual noise.

Figure 6-21 Reference model of virtual noise



For the virtual noise PSD mask to resemble the noise PSD mask in practical application, statistics on noise of the entire spectrum over a long period must be collected, as shown in Figure 6-22.

Figure 6-22 Virtual noise PSD mask



As shown in Figure 6-22, the virtual noise PSD mask more resembles the noise PSD mask than the noise margin, and ensures a more stable line and better line performance. In the meanwhile, however, virtual noise always presumes the maximum noise under the most unfavorable conditions. Therefore, line stability and low BER are achieved by compromising the connection rate.

NOTE

Figure 6-22 shows the statistical results as an example. In practical application, different carriers may use different tools and methods for collecting and analysing statistics, and the present of the statistical results may be different..

In line with ITU-T Recommendation G.997.1, the Huawei access device allows users to enable or disable virtual noise, and configure the noise margin profile and virtual noise profile by running the **xdsl noise-margin-profile add** and **xdsl virtual-noise-profile add** commands, respectively. A virtual noise profile includes multiple virtual noise PSD breakpoints. Based on this profile, the system draws the virtual noise mask for the entire spectrum using an interpolation algorithm. This process is similar to that for drawing a management information base (MIB) PSD mask.

6.4.3 Techniques for Reducing Interference

To minimize mutual interference between VDSL2 and other transmission systems, VDSL2 uses flexible mechanisms for controlling the transmit power. As these mechanisms shape the power spectral density (PSD), they are referenced as PSD shaping.

MIB-controlled PSD Mask

ITU-T Recommendation G.993.2 defines management information base (MIB)-controlled power spectral density (PSD) mask for a system to flexibly control PSD. "MIB-controlled" means configuring PSD masks through the network management system (NMS) or through a digital subscriber line access multiplexer (DSLAM). MIB-controlled PSD masks provide users with more options than the limit PSD masks defined in the standard. Carriers can control the power spectrum and reduce crosstalk by configuring suitable PSD masks according to DSLAM distribution, distance to users, and coexistence of ADSL and VDSL. Such user-configured PSD masks are referred to as MIB-controlled PSD masks.

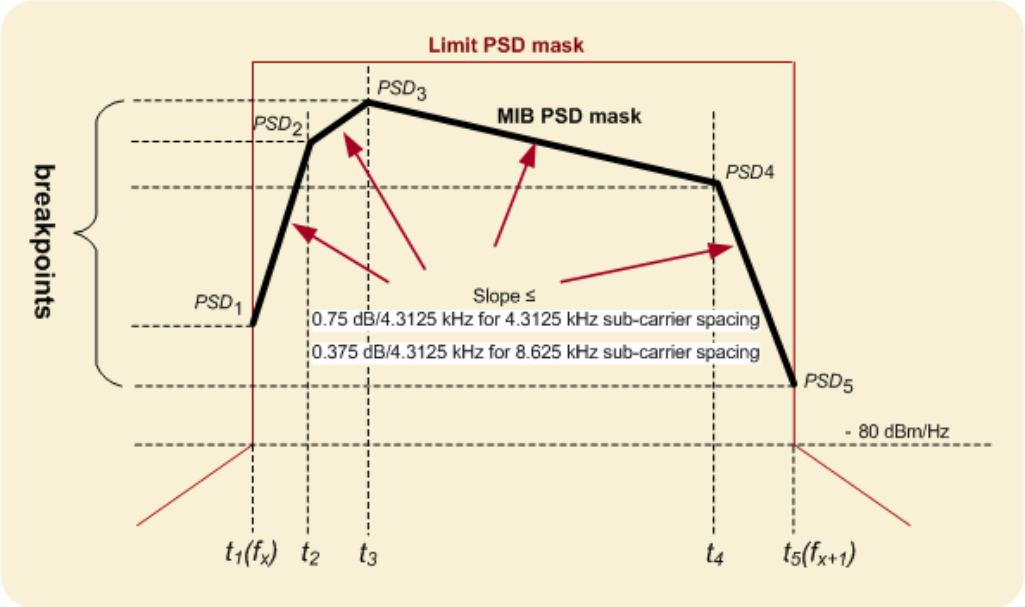
Figure 6-23 shows a common MIB-controlled PSD mask defined in ITU-T Recommendation G.993.2.

- The MIB-controlled PSD mask defines the PSD at a series of breakpoints on the transmission frequency band. Based on the PSD mask, the system determines the PSD of each subcarrier (or tone) using interpolation between two breakpoints.

- For each breakpoint, a subcarrier index (t_n) and PSD value (PSD_n) are defined. Then breakpoints are expressed like [(t_1, PSD_1), (t_2, PSD_2),…, (t_n, PSD_n)], where t_1 indicates the start frequency and t_n the stop frequency of the frequency band.
- In Figure 6-23, the limit PSD mask only indicates that the MIB-controlled PSD mask should always lie below the limit PSD mask (if the former lies above the latter, the system chooses the smaller one as the PSD mask). The turns at the PSD mask cannot form a right angle, and the slope for each turn is restricted to avoid a sharp change in the transmit power.

In addition, a maximum of 16 breakpoints can be configured in the upstream direction (for ADSL2+, a maximum of 4 breakpoints can be configured in the upstream directio) and 32 in the downstream direction. The US0 band cannot include any breakpoint.

Figure 6-23 MIB-controlled PSD mask



On the Huawei access device, users can configure MIB-controlled PSD masks by running the **xdsl mode-specific-psd-profile add** command.

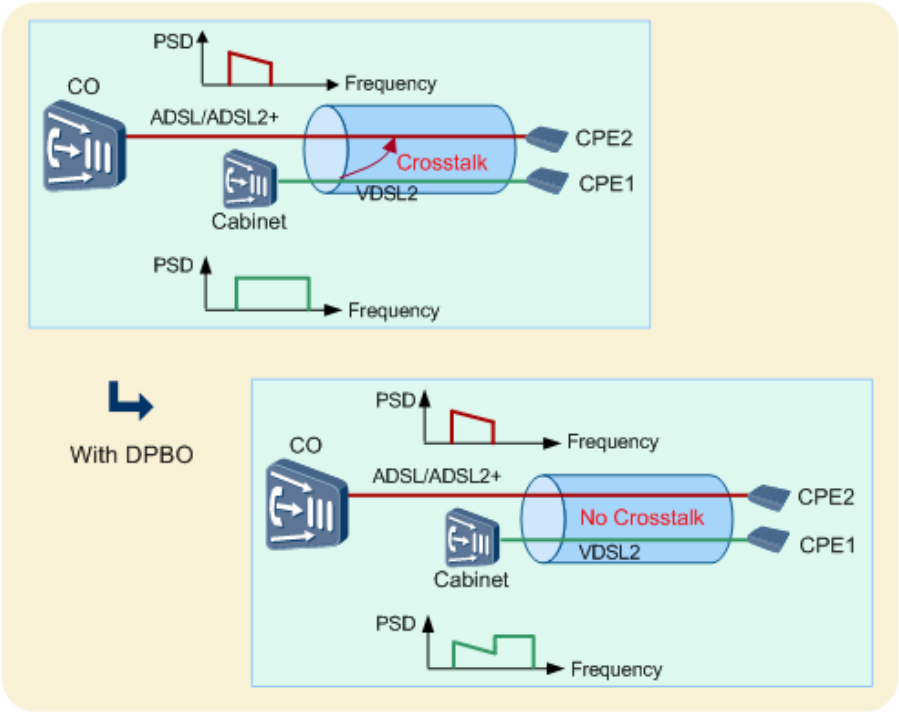
DPBO

Downstream power back-off (DPBO) is implemented to minimize crosstalk among the upstream lines in the same bundle (VDSL2 and ADSL/ADSL2+).

Definition of DPBO

On most conditions, VDSL2 lines are shorter than ADSL/ADSL2+ lines. This is why ADSL/ADSL2+ is deployed at CO and VDSL2 at cabinets, which are close to users, as shown in Figure 6-24.

Figure 6-24 Minimizing Inter-Line Crosstalk



Generally, after signals reach a cabinet, the downstream transmit power of CO is attenuated to far lower than the downstream transmit power of the cabinet. If VDSL2 and ADSL/ADSL2+ lines are deployed in the same cable bundle, the downstream signals of the cabinet have intensive crosstalk with the downstream signals of CO, which may be as intensive as to cause BER over $^{-7}$ and deteriorate services.

To minimize the inter-line crosstalk, DPBO is implemented to decrease the downstream transmit power of the cabinet so that it is close to the power of the CO-transmitted signals reaching the cabinet. Then the inter-line crosstalk is minimized.

ITU-T G.997.1 defines an algorithm for calculating DPBO, or the cabinet-end DPBO PSD mask. More specifically, the CO-end downstream PSD minus the power attenuated over the L (distance between the CO and cabinet) is equal to the PSD from the CO to cabinet. Then the cabinet-end downstream PSD is adjusted to close to the PSD.

DPBO Configuration

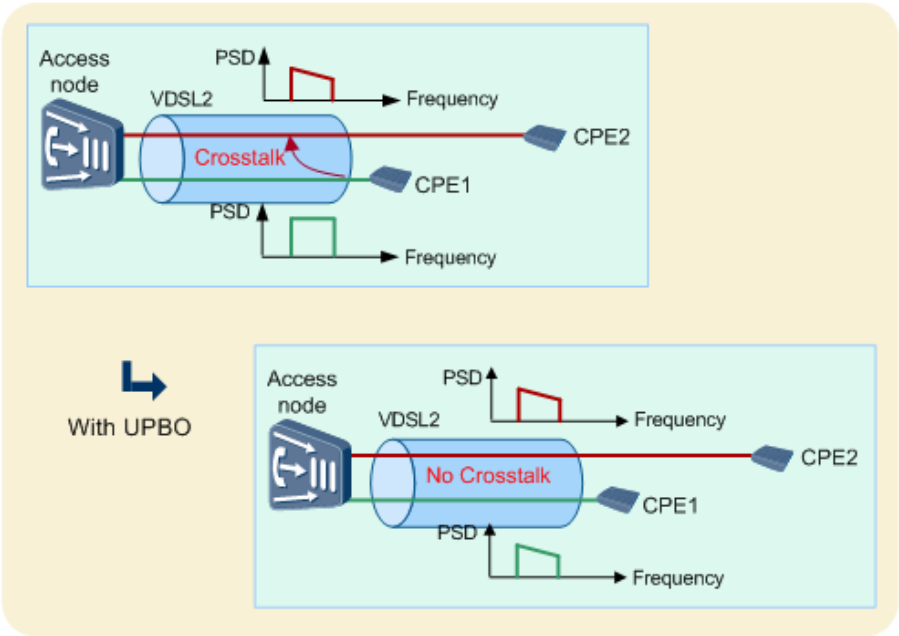
For DPBO to apply, some parameters regarding DPBO PSD mask calculation must be configured. For a Huawei access device, DPBO parameters include standard ones defined in ITU-T G.997.1, and non-standard ones customized for carriers (for ADSL2+, does not contain the non-standard ones). Users can configure DPBO by running the **xdsl dpbo-profile add** commands. For details on the parameters, see the description of the **xdsl dpbo-profile add** command.

UPBO

Upstream power back-off (UPBO) is implemented to improve spectral compatibility among VDSL2 loop systems with varied lengths and minimize crosstalk among the upstream lines.

Definition of UPBO

Figure 6-25 Minimizing Inter-line Crosstalk



As shown in Figure 6-25, VDSL2 loops deployed in the same bundle of cables may have varied lengths. The power spectral density (PSD) of the signals transmitted from CPE to CO has been severely attenuated for long VDSL2 loops, but that for short VDSL2 loops is still high. The high PSD of short VDSL2 loops will generate severe far-end crosstalk to long VDSL2 loops, impacting the upstream rate of the long loops.

VDSL2 UPBO mechanism: UPBO reduces the upstream transmit power for CPE on short VDSL2 loops while sustaining proper performance for short VDSL2 loops. In this way, signals of long and short VDSL2 loops will have similar PSDs when the signals arrive at CO, significantly reducing far-end crosstalk on long VDSL2 loops and improving their upstream transmission performance. As the upstream transmit power is reduced for short VDSL2 loops, the downstream rate of short VDSL2 loops will also decrease.

UPBO brings the following benefits:

- Minimizes the crosstalk among upstream bands for VDSL2 loops with varied lengths in a cable bundle.
- Reduces power consumption of CPE and electromagnetic radiation.

UPBO Configuration

UPBO parameters must be set for CO and CPE devices to interoperate so as to implement UPBO. For the Huawei access device, UPBO parameters include standard parameters defined in ITU-T Recommendations G.993.2 and G.997.1, and non-standard parameters customized for carriers (see Table 6-10 for details), which can be configured by running the **xdsl upbo-profile add** command.

Table 6-10 UPBO parameters

Parameter	Description	Definition	Setting
Upstream electrical length	Indicates the electrical length of the line.	• It is defined in ITU-T Recommendation G.997.1 and is similar to "Electrical length" defined in ITU-T Recommendation G.993.2. It is represented by $\mathbf{kl_0}$ in the UPBO PSD mask calculation formula. • It is equivalent to the attenuation (dB) of a given loop that has the ideal $\sqrt{f}$ attenuation feature at the 1 MHz frequency.	• This parameter refers to **Electrical length** for the **xdsl upbo-profile add** command, and it must be set when **Force CO-MIB electrical length** is set to **override**. • A carrier that requires the override mode must provide this parameter, the value of which is associated with cable specifications. It is recommended to use the auto mode, in which case this parameter does not require configuring.
Force CO-MIB electrical length	Specifies whether CPE must use the electrical length configured on CO to calculate the UPBO PSD mask.	Defined in ITU-T Recommendation G.997.1, it indicates how CPE obtains $\mathbf{kl_0}$.	This parameter refers to **Will you force the CPE to use the electrical length to compute the UPBO** in the **xdsl upbo-profile add** command. Based on the options provided in the standard, the following values are designed for this parameter: • auto: optional. CPE selects a proper way of obtaining $\mathbf{kl_0}$. The following ways are available for CPE: – 1-max(kl0_CO,kl0_CPE): The greater one of the $\mathbf{kl_0}$ values calculated by CO and CPE applies. – 2-min(kl0_CO,kl0_CPE): The smaller one of the $\mathbf{kl_0}$ values calculated by CO and CPE applies. – 3-kl0_CO: The $\mathbf{kl_0}$ value calculated by CO applies. – 4-kl0_CPE: The $\mathbf{kl_0}$ value calculated by CPE applies. These four ways are carrier-customized and beyond the scope of the standard. Carriers will choose a proper way for CPE to obtain $\mathbf{kl_0}$. If carriers do not choose one, **2-min(kl0_CO,kl0_CPE)** is recommended. In addition, the selected way applies only when **UPBO electrical length**

Parameter	Description	Definition	Setting
			estimation mode is set to **0-ELE_M0**. • override: mandatory. CPE must use $\mathbf{kl_0}$ configured on CO (or the above-mentioned **Electrical length** parameter). • disableUPBO: UPBO is disabled.
UPBO reference PSD per band	Calculates **UPBOPSD** for the upstream and downstream bands (except US0).	**UPBOPSD** is a key parameter in the formula for calculating UPBO PSD mask and includes two sub-parameters: • Sub-parameter a: ranges from 40 dBm/Hz to 80.95 dBm/Hz and changes at a step of 0.01 dBm/Hz. • Sub-parameter b: ranges from 0 dBm/Hz to 40.95 dBm/Hz and changes at a step of 0.01 dBm/Hz.	• This parameter refers to **UPBO reference PSD per band** for the **xdsl upbo-profile add** command, and sub-parameters a and b need to be set for different upstream bands. • The values of sub-parameters a and b vary according to regions. Some Annex appendixes in ITU-T Recommendation G.993.2 and region-specific standards, such as T1.417, ETSI101388, and ETSI101271 define reference values and calculation methods for the two sub-parameters. Generally, carriers specify values for the two sub-parameters.
UPBO reference electrical length per band	Indicates $\mathbf{kl_{0\_REF}}$ of each upstream band (except US0).	ITU-T Recommendation G.993.2 defines the following methods for calculating the UPBO PSD mask: • Reference PSD UPBO method (mandatory) • Equalized FEXT UPBO method (optional): The calculation includes a far-end crosstalk factor and therefore is more accurate. The device, regardless of its supplier, must support the first method. The second method is optional and is not supported by some CPEs. When the second method is used, $\mathbf{kl_{0\_REF}}$ is required, which refers to the far-end crosstalk factor.$\mathbf{kl_{0\_REF}}$ ranges from 1.8 dB to 63.5 dB and changes at a step of 0.1 dB. The value **0** indicates not	• This parameter refers to **UPBO reference electrical length per band** of the **xdsl upbo-profile add** command, and it needs to be set for different upstream bands. • Generally, carriers configure the parameter value. If carriers do not configure the parameter value, refer to the reference values and calculation methods defined in ITU-T Recommendation G.993.2 and region-specific standards, such as T1.417, ETSI101388, and ETSI101271.

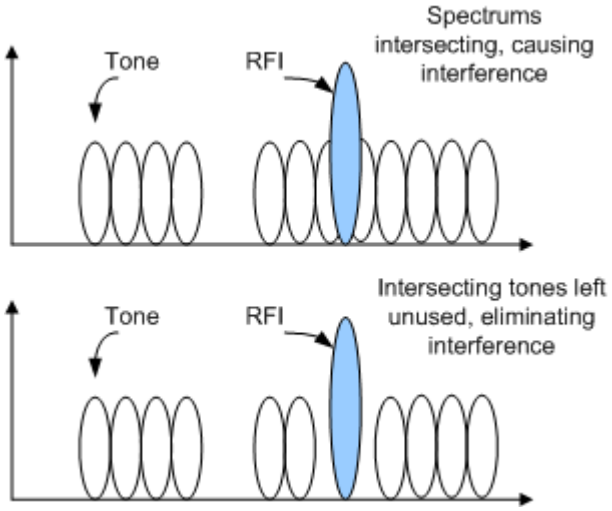
Parameter	Description	Definition	Setting
		using the far-end crosstalk factor.	
UPBO electrical length estimation mode	Indicates the mode for estimating the UPBO electrical length.	• This parameter is defined in ITU-T Recommendation G.997.1 and refers to **Electrical Length Estimation Method** defined in ITU-T Recommendation G.993.2. It indicates how CO and CPE estimate $\mathbf{kl_0}$. • ITU-T Recommendation G.993.2 defines the following modes of estimating $\mathbf{kl_0}$: – 0-ELE_M0 – 1-ELE_DS – 2-ELE_PB – 3-ELE_MIN • When **Force CO-MIB electrical length** is set to **auto**, CO and CPE estimate $\mathbf{kl_0}$ using the calculation method specified by this parameter.	• This parameter refers to **UPBO electrical length estimation mode** of the **xdsl upbo-profile add** command and is generally specified by carriers. • This parameter has a lower priority than **Force CO-MIB electrical length**. In other words, the parameter setting applies to UPBO PSD mask calculation only when **Force CO-MIB electrical length** is set to **auto**.
UPBO electrical length threshold percentile	Indicates the minimum threshold percentile of the UPBO electrical length.	• This parameter is defined in ITU-T Recommendation G.997.1 and refers to **UPBO Electrical Length Minimum Threshold (UPBOELMT)** in the ITU-T Recommendation G.993.2-defined UPBO PSD mask calculation formula. • This parameter will be used in the UPBO PSD mask calculation formula only when **UPBO electrical length estimation mode** is set to a mode other than **ELE_M0**.	• This parameter refers to **UPBO electrical length threshold percentile** of the **xdsl upbo-profile add** command and is generally specified by carriers or set to default. • The parameter setting applies to UPBO PSD mask calculation only when **UPBO electrical length estimation mode** is set to a mode other than **ELE_M0**.
UPBO Boost Mode	Enables or disables forcible correction of	• Not all devices support the calculation that includes the far-end crosstalk factor. Though CPE does	This parameter refers to **UPBO Boost Mode** of the **xdsl upbo-profile add** command. Set this parameter according to carriers'

Parameter	Description	Definition	Setting
	kl_0.	not support the far-end crosstalk factor, some carriers may require the far-end crosstalk factor to be effective. To address this requirement, correction of **kl_0** can be enabled. Then CO sends the corrected **kl_0** calculation formula to CPE in order to forcibly correct **kl_0** estimated by CPE, achieving similar calculation including the far-end crosstalk factor. • This parameter is not a standard parameter.	requirements.

RFI Notching

VDSL2 uses a wide range of frequencies, with the highest frequency of 30 MHz, which covers the medium wave, short wave, and ham radio. Therefore, VDSL2 has to provide a solution to radio frequency interference (RFI). There are complex RFI factors, and the conventional countermeasures against RFI are not cost-effective. RFI lasts long and has such a narrow interference band that it is densely populated on one or several tones. RFI notching is introduced to resolve the issue.

Figure 6-26 Working principle of RFI notching



RFI notching means leaving some RFI-free tones unused to counteract RFI. Though RFI notching sacrifices some line transmission rate, it is effective. When the tones are left unused, the transmit PSD will be decreased to below the ITU-T Recommendation G.993.2-defined -80

dBm/Hz but not to none. If the tones can still carry bits with the transmit PSD below -80 dBm/Hz, the tones will carry some bits. This is how RFI notching differs from tone blackout.

In practical application, if the RFI power is intensive (no specific benchmark for the intensity), RFI notching may fail to eliminate RFI. In this case, tone blackout can black out the interference-suffering tones to avoid RFI.

On the Huawei access device, users can run the **xdsl rfi-profile add** command to configure RFI notching. The RFI notching band cannot be over-extensive or include the pilot tone; otherwise, the line may fail to be activated.

NOTE

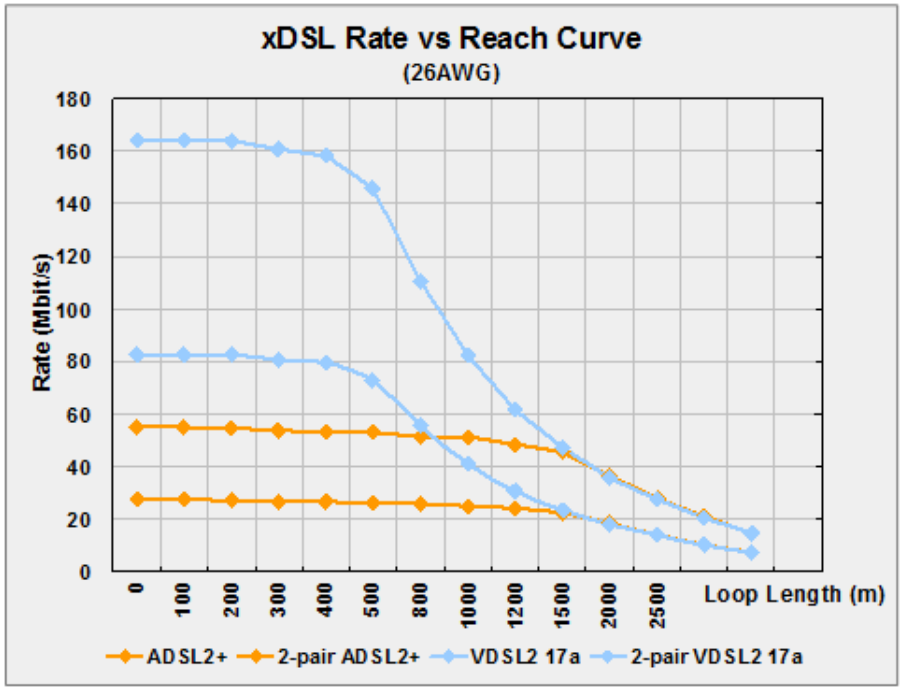
The system determines the pilot tones in line with ITU-T Recommendation G.994.1. Users can identify the pilot tones by comparing the spectrum profile against the ITU-T Recommendation G.994.1. Generally, the RFI notching band has a high frequency while the pilot tone has a low frequency, and they are less likely to intersect.

6.4.4 VDSL2 PTM Bonding

VDSL2 packet transfer mode (PTM) bonding, or VDSL2 Ethernet in the first mile (EFM) bonding, is implemented in line with ITU-T Recommendation G.998.2. It extends the access distance while maintaining a constant access rate or increases the access rate while maintaining a constant access distance, by means of bonding.

Figure 6-27 shows a comparison of rate-to-distance curves with and without bonding (based on 26AWG twisted pairs, under lab conditions).

Figure 6-27 Rate-to-distance curves (with and without bonding, taking 2-pair bonding as an example)



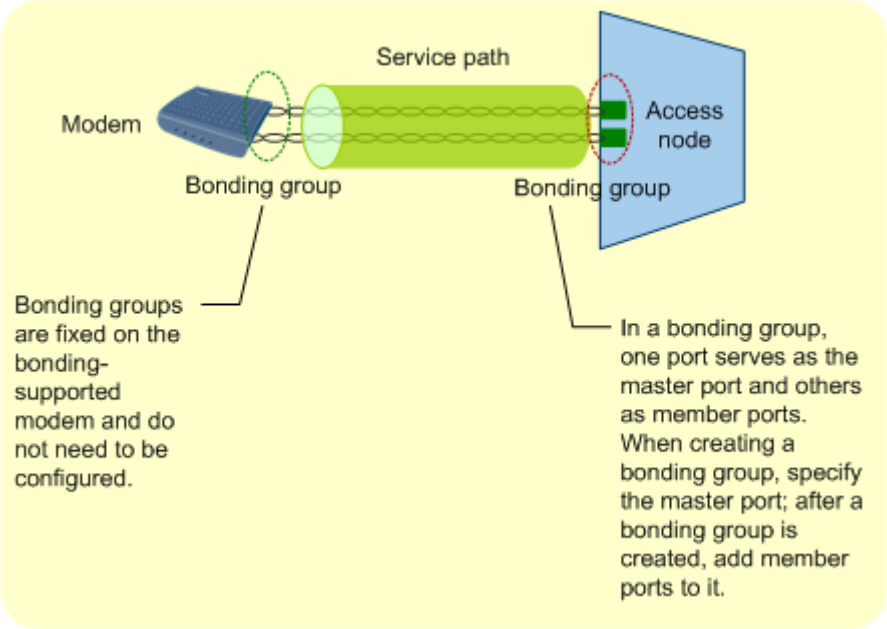
When VDSL2 PTM bonding is configured, CO divides one Ethernet packet into multiple fragments and distributes them over multiple lines leading to CPE. CPE then assembles the received fragments. The system runs the IEEE 802.3ah protocol to divide Ethernet packets

and distribute fragments. During bonding initialization, CO and CPE run the ITU-T Recommendation G.994.1 to negotiate on bonding.

VDSL2 PTM Bonding Configuration

Bonded VDSL2 ports form a bonding group, one serving as the master port and others as member ports, as shown in the following figure. Services can be configured only on the master port in a bonding group.

Figure 6-28 Application of VDSL2 PTM bonding (taking 2-pair bonding as an example)

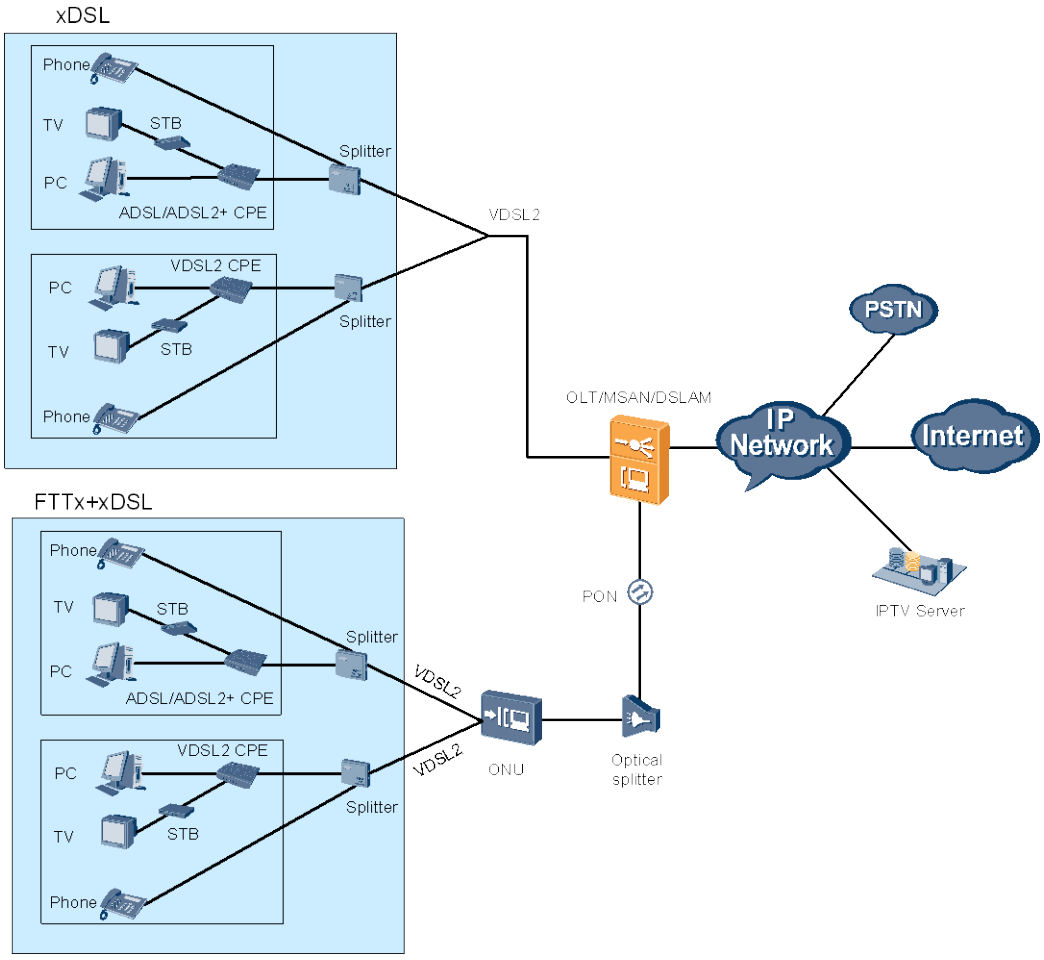


6.5 VDSL2 Deployment and Maintenance

6.5.1 VDSL2 Network Applications

This topic describes the network applications of the VDSL2 access feature.

Figure 6-29 VDSL2 network applications



As shown in the figure above, typical scenarios of VDSL2 network application are as follows.

1. The MA5600T/MA5603T/MA5608T directly provides the VDSL2 access.

   On the user side, VDSL2 CPEs (working in the PTM mode) or ADSL/ADSL2+ CPEs (working in the ATM mode) can be connected to the MA5600T/MA5603T/MA5608T to provide high-speed Internet access service, video service and public switched telephone network (PSTN) voice service for subscribers.

2. The MA5600T/MA5603T/MA5608T provides PON optical ports for connecting to ONUs and the ONUs provide the VDSL2 access.

   The ONUs are placed on street side or in corridors. In the downstream direction, the ONUs provide the VDSL2 access for subscribers; in the upstream direction, the ONUs are connected to the MA5600T/MA5603T/MA5608T by PON. The FTTx+VDSL2 network topology addresses the distance restriction on the VDSL2 access.

6.5.2 VDSL2 Engineering Precautions

The quality of the DSL feature depends on the line quality. Take the following precautions when deploying the VDSL2 feature.

1. It is recommended that the line distance be smaller than 1000 meters.

VDSL2 requires high working frequency. For long distance transmission, the attenuation is large and the high frequency data traffic decreases. After a distance of 1.2 km, VDSL2 has similar performance as ADSL2+.

2. It is recommended that the diameter of twisted pairs be 0.4 mm or larger.

   For a cable with a certain length, a smaller cable diameter results in a larger loop resistance and signal attenuation. In some projects, parallel cables are used as drop cables. This is not standard because it will cause many issues, such as reduced line activation rates and even line activation failures. Therefore, ensure that standard twisted pairs are used as drop cables.

3. A protective unit using a gas discharge tube is recommended for the main distribution frame (MDF).

4. ADSL2+ splitters cannot be used at the user end. Instead, ADSL2+/VDSL2 compatible splitters or dedicated VDSL2 splitters must be used.

   ADSL2+ splitters cannot meet the requirements of VDSL2 in terms of the frequency response and line longitudinal balance in the high frequency. These two indicators determine the performance of the VDSL2 feature and may result in a failure in VDSL2 line activation or frequent disconnections in worst cases.

5. The insertion attenuation between the wiring terminal and fiber distribution terminal (FDT) is small. Even so, make sure that they are in good contact, cables are routed properly and connected securely, and wiring terminals are in good condition, to prevent unexpected signal attenuation and crosstalk and therefore to ensure the stability of the VDSL2 feature.

6. Prevent line aging caused by factors such as line exposure.

7. Avoid bridge taps in the subscriber line loop.

   A bridge tap is an idle twisted pair with one end connecting to the trunk cable or FDT and the other end open. It is usually used to ensure the flexibility of subscriber line loops. A bridge tap results in resistance mismatch. Signal reflection occurs at a bridge tap and therefore the signal attenuation is very large. This greatly affects the activation rate of VDSL2.

   In an actual project, a cable is used to connect the subscriber splitter and telephone terminal. If the cable does not connect to a telephone terminal, it is called a bridge tap. The impact of a bridge tap on the VDSL2 upstream and downstream rates increases with the length of the bridge tap.

8. No telecommunication devices are configured between the splitter and the drop cable connected to the user,

   including fax machines, phone extensions, IP dialers, audio modems, and anti-theft devices. These devices can only be connected to the telephone outlet of the splitter. If multiple voice devices are configured between the splitter and the drop cable connected to the user, a splitter must be configured in front of each voice device.

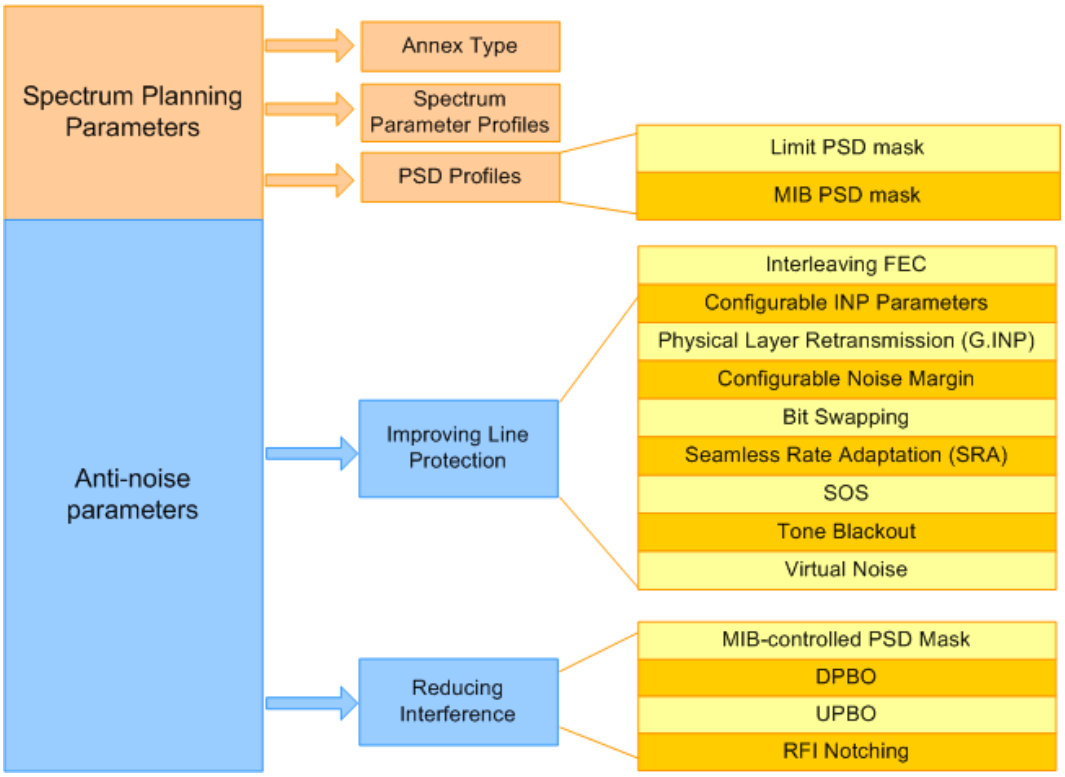
9. The drop cable connected to the user must stay away from household appliances, such as air conditioner outdoor units, refrigerators, and sound boxes. Otherwise, industrial frequency noise will increase.

6.5.3 Brief Introduction to VDSL2 Configurations and Applications

This section describes roadmap for VDSL2 configurations and applications.

VDSL2 line configuration involves two types of important parameters, shown in Figure 6-30.

Figure 6-30 Diagram for VDSL2 line parameter configuration



1. Set spectrum planning parameters (for details, see 6.3.1 Overview of VDSL2 Spectrum Planning).
   a. Choose an appropriate transmission mode (that is, the applied standard and 6.3.2 Annex Types and US/DS Frequency Band Planning) depending on the DSL network plan and deployment.
   b. Choose a 6.3.5 Spectrum Parameter Profiles (8 in total, 8a-30a) depending on requirements for spectrum parameters.
   c. Configure 6.3.6 PSD Profiles based on power spectrum requirements. (You can choose an Annex-defined 6.3.7 Limit PSD Mask or manually configure a 6.3.9 MIB PSD Mask.)
2. Set anti-noise parameters to achieve a balance between performance and reliability (for details, see 6.4.2 Key Techniques for Improving Line Protection and 6.4.3 Techniques for Reducing Interference).

   Various noise interferences exist on a subscriber digital line. VDSL2 provides a number of countermeasures to improve line stability, achieving higher line quality, and a lower packet loss ratio and bit error ratio. In most cases, stability is improved at the expense of line performance, for example, by reducing the activation rate or prolonging service latency. It is necessary, therefore, to set appropriate line parameters to achieve a balance between line reliability and performance. Table 6-12 lists the impact of various noise-cancellation countermeasures on line performance.

Table 6-11 Impact of countermeasures on line performance

Category	Countermeasure	Activation Rate Affected or Not	Service Latency Prolonged or Not

Category	Countermeasure	Activation Rate Affected or Not	Service Latency Prolonged or Not
Improving line protection capabilities (passive defense against noise interference)	Interleaving FEC	Yes	Yes
	Configurable INP Parameters	Yes	Yes
	Physical Layer Retransmission (G.INP)	Yes	Yes
	Configurable Noise Margin	Yes	No
	Bit Swapping	No	No
	SRA	No; the line rate is dynamically adjusted after a line is activated.	Yes (SRA may change the interleaving depth, resulting in latency deviations.)
	SOS	No; the line rate is dynamically adjusted after a line is activated.	Yes (SRA is usually required for the use of SOS and service latency will be prolonged.)
	Tone Blackout	Yes	No
	Virtual Noise	Yes	No
Reducing interference output These countermeasures mitigate the impact of a line on other transmission systems. To achieve this, noise interference on the line must be reduced, mainly using power spectrum density (PSD) shaping	MIB-controlled PSD Mask	Yes	No
	DPBO	Yes	No
	UPBO	Yes	No
	RFI Notching	Yes	No

Table 6-12 lists techniques that counter different types of noises.

Table 6-12 Types of noises and countermeasures

Noise Type	Noise Characteristics	Countermeasure	Description

Noise Type	Noise Characteristics	Countermeasure	Description
Pulse noises	Pulse noises are intensive, brief (micro- or milliseconds), and cover the entire frequency band. Pulse noise may derive from on-hook/off-hook of telephones, power-on/power-off of home appliances, or natural electricity discharge.	• Interleaving FEC • Configurable INP Parameters • Physical Layer Retransmission (G.INP)	• Interleaving FEC, when used with erasure decoding, significantly improves system noise resistance. • To help users select appropriate INP parameter values during configuration, VDSL2 introduces the impulse noise monitoring (INM) technique. For details on erasure decoding and INM, see Configurable INP Parameters.
Environmental noises, such as background noise and noise caused by changes in temperature or relative humidity levels.	Noise that lasts a long period of time (microseconds), covers a narrow spectrum range, has a weak intensity, and changes slowly. Such a noise may come from amateur radio interference (such as that generated by remotely-controlled toys) and may overlap with radio frequency interference (RFI) described below.	Bit Swapping	In ITU-T Recommendation G.993.2, bit swapping, SRA, and SOS are on-line reconfiguration (OLR) techniques.
	Noise that lasts a long period of time (seconds), covers a wide spectrum range, has a weak intensity, and changes slowly.	SRA	
	Noise that lasts a long period of time (seconds), covers a wide spectrum range, has a strong intensity, and changes fast.	SOS	
	Noise that lasts a long period of time (seconds), covers a wide spectrum range, and has a constant intensity.	• Configurable Noise Margin • Virtual Noise	The Configurable Noise Margin technique is widely used.
RFI	RFI noise covers a narrow spectrum range, and interference occurs mostly on one or more tones. This type of noise mainly derives from broadcast and	• RFI Notching • Tone Blackout • Bit Swapping	The RFI Notching technique is recommended.

Noise Type	Noise Characteristics	Countermeasure	Description
	amateur radio communication.		
Inter-line crosstalk	Inter-line crosstalk refers to the noise caused by crosstalk between lines in a bundle, and it is associated with distribution of DSLAMs, distance to users, and coexistence of ADSL and VDSL2.	• DPBO • UPBO • MIB-controlled PSD Mask • Bit Swapping • SOS • Configurable Noise Margin • Virtual Noise	The DPBO or UPBO technique is recommended.

6.5.4 Configuring VDSL2 Access

VDSL2 service configuration includes VDSL2 profile configuration and VDSL2 user port configuration. This topic describes the detailed configuration methods and procedures.

Overview of Configuring VDSL2 Templates and Profiles

As mentioned in *Brief Introduction to VDSL2 Configurations and Applications*, spectrum parameter and anti-noise parameter configurations are the key points in VDSL2 line parameter configuration. Spectrum and anti-noise parameters are configured in a VDSL2 line parameter profile. In addition to the line parameter profile, the VDSL2 alarm template can be configured to facilitate line maintenance. After the line parameter profile and alarm template are configured, they can be used for activating DSL ports.

Context

The device supports three VDSL2 modes: normal (TR129), TI, and TR165. Run the **switch vdsl mode to** to switch between the modes. By default, the TR129 mode is used.

The alarm template configuration is the same for the three modes but the line parameter profile configuration varies with the VDSL2 mode.

Configuring a VDSL2 Alarm Template

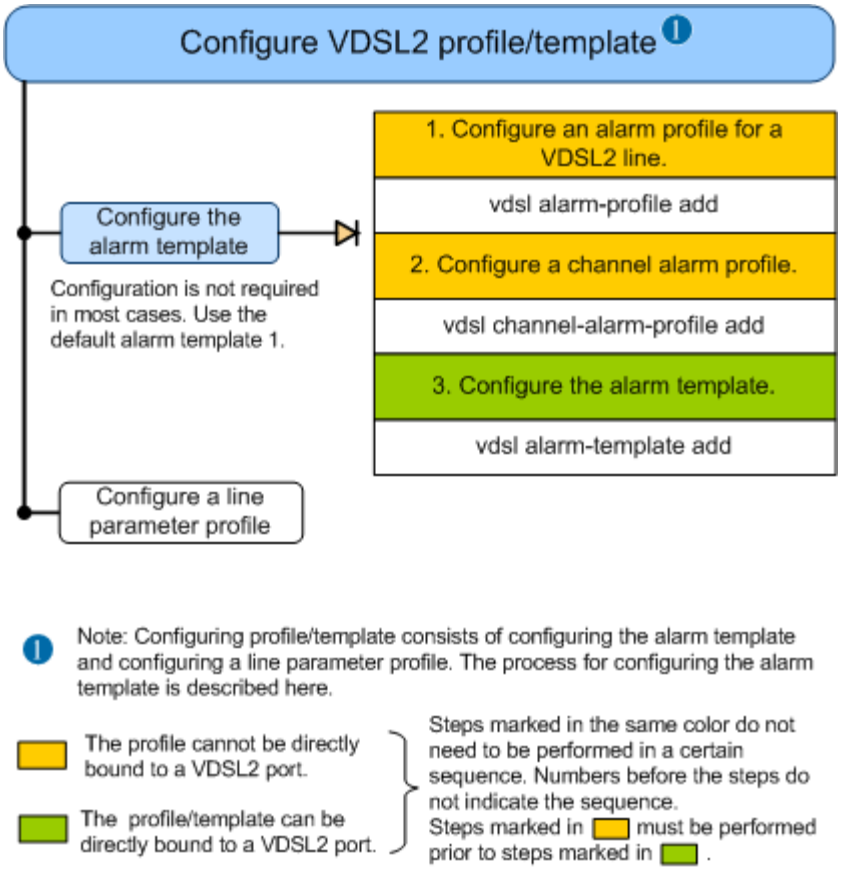
A VDSL2 alarm template that is used for activating ports consists of a line alarm profile and a channel alarm profile.

Context

In most cases, there is no need to configure a VDSL2 alarm template. You can use the default alarm template 1.

If you want to configure the VDSL2 alarm template, follow the process described in Figure 6-31.

Figure 6-31 Flowchart for configuring a VDSL2 alarm template



Procedure

Configure a VDSL2 line alarm profile.

Run the **vdsl alarm-profile quickadd** command to quickly add a VDSL2 line alarm profile, or run the interactive command **vdsl alarm-profile add** to add a VDSL2 line alarm profile.

Step 1 Configure a VDSL2 channel alarm profile.

Run the **vdsl channel-alarm-profile quickadd** command to quickly add a VDSL2 channel alarm profile, or run the interactive command **vdsl channel-alarm-profile add** to add a VDSL2 channel alarm profile.

Step 2 Configure a VDSL2 alarm template.

Run the **vdsl alarm-template quickadd** command to quickly add a VDSL2 alarm template, or run the interactive command **vdsl alarm-template add** to add a VDSL2 alarm template.

The main parameters are as follows:

- **line** *alarm-profile-index*: indicates the line alarm profile in the alarm template. If this parameter is required, configure it prior to **channel1**.
- **channel1** *channel1-alarm-profile-index*: indicates the channel alarm profile for channel 1 in the alarm template.

- **channel1** *channel1-alarm-profile-index*: indicates the channel alarm profile for channel 2 in the alarm template. Channel 2 is unavailable and this configuration will not take effect. Therefore, there is no need to set this parameter.

Step 3 Check if the configurations in the alarm template agree with the data plan.

Run the **display vdsl alarm-template** command to check if the configurations in the alarm template agree with the data plan.

After the alarm template is successfully configured, it can be directly used for activating VDSL2 ports.

----End

Example

To add alarm template 3 that uses channel alarm profile 1 (default) and line alarm profile 2 with alarming upon receiving error sample packets disabled, do as follows:

```
huawei(config)#vdsl alarm-profile quickadd 2 received-ES-abnormal-alarm disable
huawei(config)#vdsl alarm-template quickadd 3 line 2 channel1 1
huawei(config)#display vdsl alarm-template 3
```

Configuring a VDSL2 Line Parameter Profile

A VDSL2 line parameter profile is the key for VDSL2 service configurations. This topic describes how to configure the VDSL2 line parameter profiles in different VDSL2 modes.

Prerequisites

Run the **display xdsl mode** command to check whether the VDSL2 mode is the desired mode. The default mode is TR129.

If the current mode is not the desired one, run the **switch vdsl mode to** command in diagnose mode to switch the mode to the desired mode.

NOTE
When both the ADSL2+ and VDSL2 modes are TR165, the configured profile is used by both ADSL2+ and VDSL2 ports. If only one of the ADSL2+ and VDSL2 modes is TR165, the configured profile is used only by the one in TR165 mode.

Typical Configuration Reference

Key parameters to be configured do not vary with the VDSL2 mode but belong to different profiles. Hence, commands for configuring the key parameters vary with the VDSL2 mode.

Table 6-13 lists the typical configurations for key parameters of a VDSL2 line (26 AWG, 0.4 mm twisted pair). Information provided in this table is for reference only. Usually, take the default values for the other parameters.

Table 6-13 Typical configurations for key parameters of a VDSL2 line (0.4 mm twisted pair)

Parameter	Access Distance				Remarks
	< 300 m	300–500 m	500–800 m	800–1000 m	
Selected 6.3.5 Spectrum	17a	12a	8b, 12a	8b	-

Parameter	Access Distance				Remarks
	< 300 m	300–500 m	500–800 m	800–1000 m	
Parameter Profiles					
6.3.7 Limit PSD Mask	B8-11 (998ADE17-M2x-A)	B8-6 (998-M2x-B)	B8-6 (998-M2x-B)	B8-6 (998-M2x-B)	-
Enable 6.3.2 Annex Types and US/DS Frequency Band Planning (U0)	Yes	Yes	Yes	Yes	Enable US0 if the distance is longer than 500 m.
Maximum transmit rate downstream	50 Mbit/s	40 Mbit/s	25 Mbit/s	20 Mbit/s	• Limiting the upstream and downstream rates ensures a higher signal-to-noise ratio (SNR) margin for the line and therefore enhances its capability for withstanding noise and interference. • The rates can be specified using these two parameters and also can be specified in the traffic profile. When they are specified in both ways, the actual activation rate is determined by the smaller one.
Maximum transmit rate upstream	15 Mbit/s	10 Mbit/s	5 Mbit/s	2 Mbit/s	
Configurable INP Parameters	2 symbols	2 symbols	2 symbols	2 symbols	-
Configurable Noise Margin	8 dB	8 dB	8 dB	8 dB	-
Path mode	PTM	PTM	PTM	PTM	If a line is activated in

Parameter	Access Distance				Remarks
	< 300 m	300–500 m	500–800 m	800–1000 m	
					VDSL2 mode, the path mode can only be PTM; if a line is activated in ADSL, ADSL2, or ADSL2+ mode, the path mode can only be ATM. If the configured path mode is inconsistent with the mode supported by the actual physical line, the board adapts the mode to the one it supports to guarantee successful line activation first. The specified path mode is the same as the actual mode used in the activation of a VDSL2 line. Hence, **it does not need to be set during data configuration**.

Configuration Process

Figure 6-32 shows the process for configuring a VDSL2 line parameter profile in TR129 mode.

Figure 6-32 Flowchart for configuring a VDSL2 line parameter profile - TR129 mode

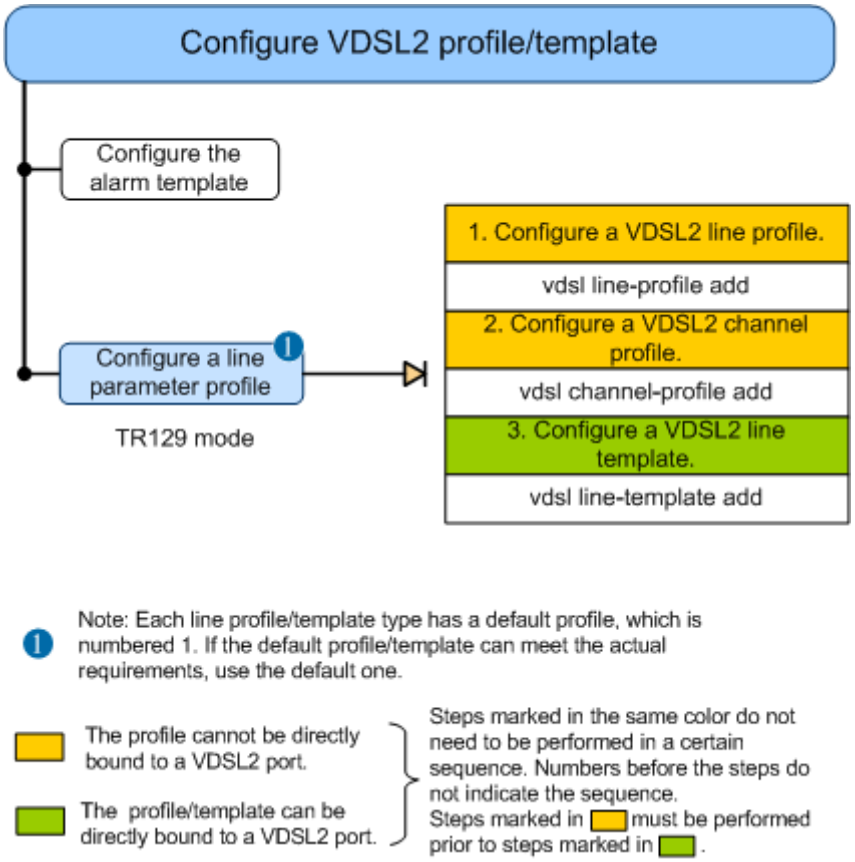


Figure 6-33 shows the process for configuring a VDSL2 line parameter profile in TI mode.

Figure 6-33 Flowchart for configuring a VDSL2 line parameter profile - TI mode

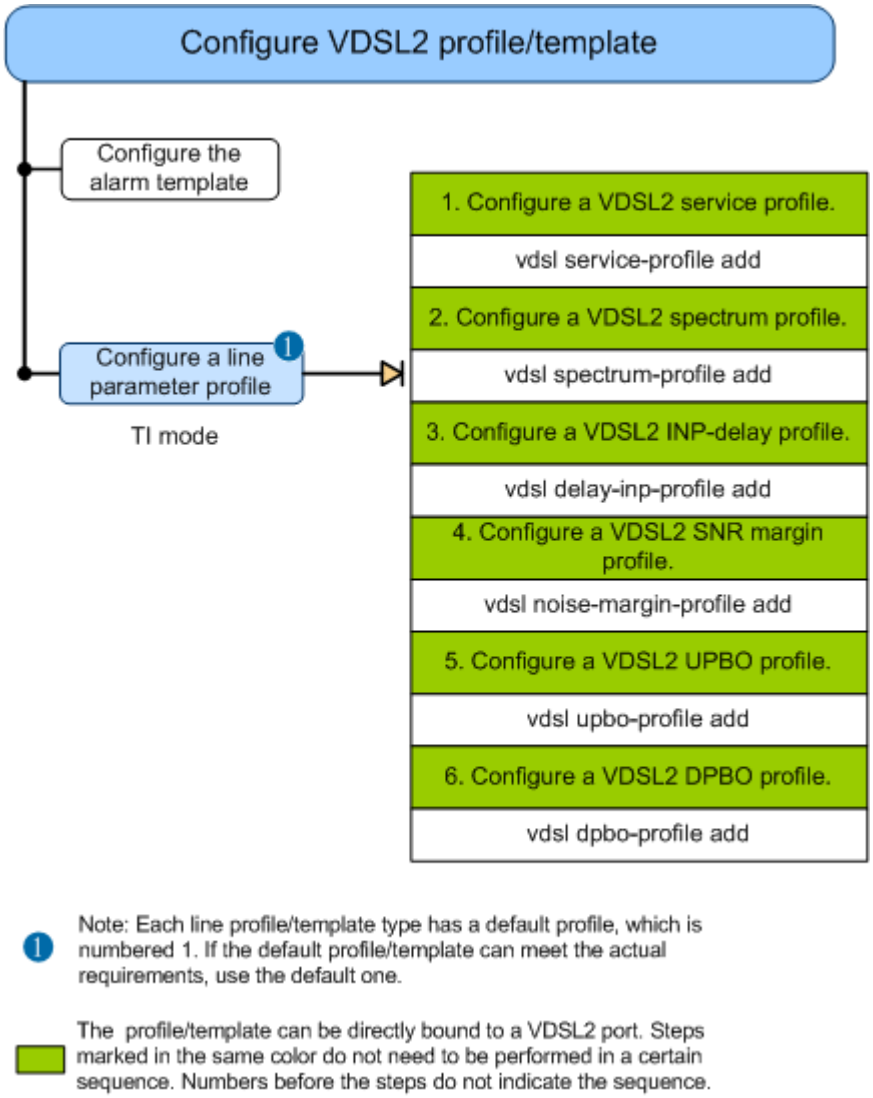
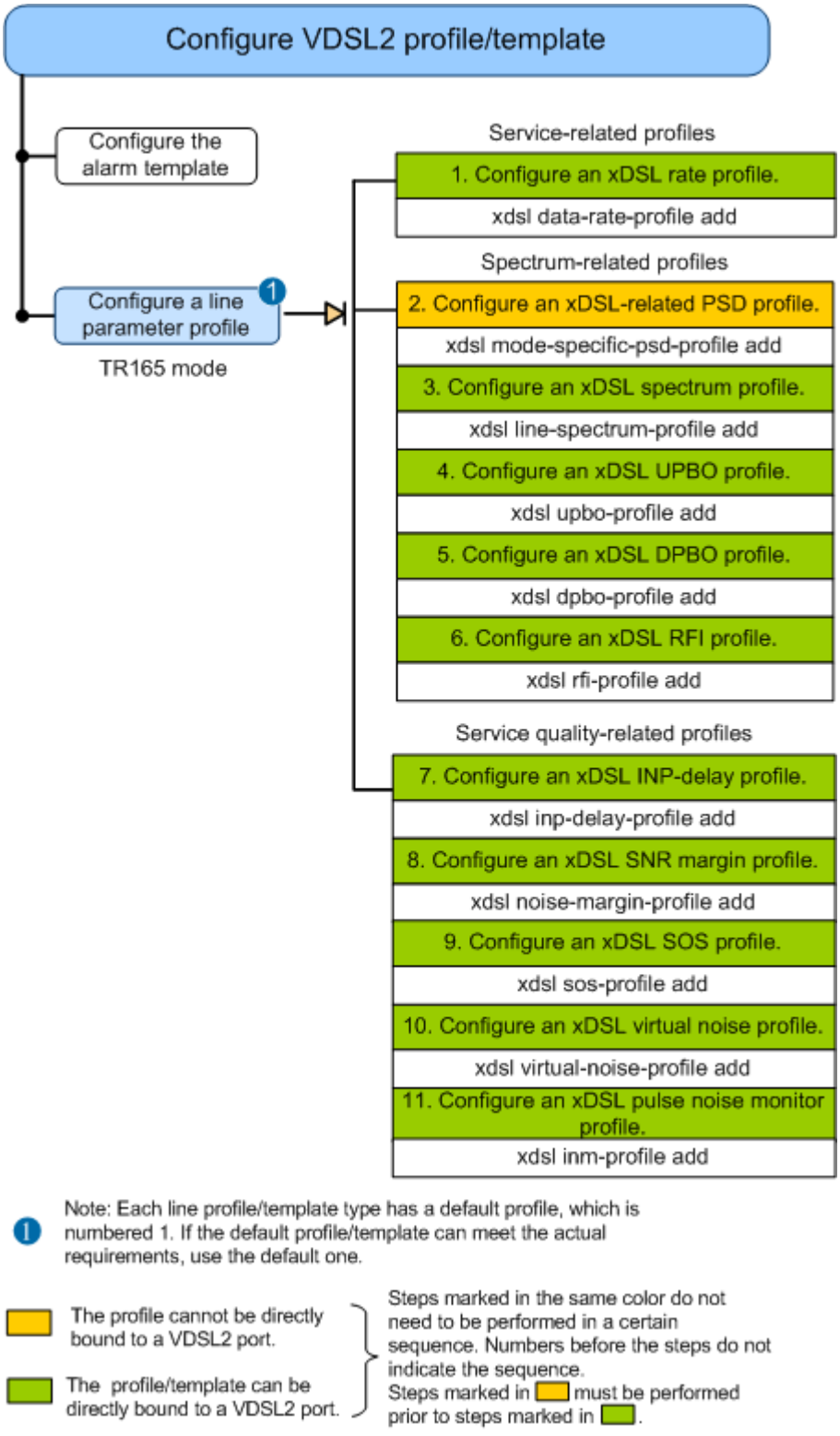


Figure 6-34 shows the process for configuring a VDSL2 line parameter profile in TR165 mode.

Figure 6-34 Flowchart for configuring a VDSL2 line parameter profile - TR165 mode



Procedure

- Do as follows to configure a VDSL2 line parameter profile when the VDSL2 mode is TR129:
  a. Configure a VDSL2 line profile.

     Run the **vdsl line-profile quickadd** command to quickly add a VDSL2 line profile; or run the interactive command **vdsl line-profile add** to add a VDSL2 line profile.
  b. Configure a VDSL2 channel profile.

     Run the **vdsl channel-profile quickadd** command to quickly add a VDSL2 channel profile, or run the interactive command **vdsl channel-profile add** to add a VDSL2 channel profile.
  c. Configure a VDSL2 line template.

     Run the **vdsl line-template quickadd** command to quickly add a line template; or run the interactive command **vdsl line-template add** to add a line template.

     The line template binds the line profile and channel profile. Only the line template is used to activate VDSL2 ports.
- Do as follows to configure a VDSL2 line parameter profile when the VDSL2 mode is TI:
  a. Configure a VDSL2 service profile.

     Run the **vdsl service-profile quickadd** command to quickly add a VDSL2 service profile, or run the interactive command **vdsl service-profile add** to add a VDSL2 service profile.
  b. Configure a VDSL2 spectrum profile.

     Run the **vdsl spectrum-profile quickadd** command to add a VDSL2 spectrum profile, or run the interactive command **vdsl spectrum-profile add** to add a VDSL2 spectrum profile.
  c. Configure a VDSL2 INP-delay profile.

     Run the **vdsl delay-inp-profile quickadd** command to add a VDSL2 INP-delay profile, or run the interactive command **vdsl delay-inp-profile add** to add a VDSL2 INP-delay profile.
  d. Configure a VDSL2 SNR margin profile.

     Run the **vdsl noise-margin-profile quickadd** command to add a VDSL2 SNR margin profile, or run the interactive command **vdsl noise-margin-profile add** to add a VDSL2 SNR margin profile.
  e. Configure a VDSL2 UPBO profile.

     Run the **vdsl upbo-profile quickadd** command to quickly add a VDSL2 UPBO profile, or run the interactive command **vdsl upbo-profile add** to add a VDSL2 UPBO profile.
  f. Configure a VDSL2 DPBO profile.

     Run the **vdsl dpbo-profile quickadd** command to quickly add a VDSL2 DPBO profile, or run the interactive command **vdsl dpbo-profile add** to add a VDSL2 DPBO profile.

  After a profile is successfully configured, it can be used for activating VDSL2 ports.
- Do as follows to configure a VDSL2 line parameter profile when the VDSL2 mode is TR165:
  a. Configure service-related profiles.
     i. Run the **xdsl data-rate-profile quickadd** command to quickly add an xDSL rate profile, or run the interactive command **xdsl data-rate-profile add** to add an xDSL rate profile.

NOTE

- When VDSL2 ports are activated in TR165 mode, the upstream rate profile and downstream rate profile are used separately. The two profiles can be one profile. However, they are usually two different profiles because the upstream and downstream rates are different in actual practice.
- It is recommended that the **Data path mode** parameter in this command take the default value. If this parameter does not take the default value, ensure that it has the same value in the upstream and downstream rate profiles that are used for activating a VDSL2 port.

b. Configure spectrum-related profiles.

i. Run the **xdsl mode-specific-psd-profile quickadd** command to quickly add an xDSL-related PSD profile, or run the interactive command **xdsl mode-specific-psd-profile add** to add an xDSL-related PSD profile.

ii. Run the **xdsl line-spectrum-profile quickadd** command to quickly add an xDSL spectrum profile, or run the interactive command **xdsl line-spectrum-profile add** to add an xDSL spectrum profile.

iii. Run the **xdsl upbo-profile quickadd** command to quickly add an xDSL UPBO profile, or run the interactive command **xdsl upbo-profile add** to add an xDSL UPBO profile.

iv. Run the **xdsl dpbo-profile quickadd** command to quickly add an xDSL DPBO profile, or run the interactive command **xdsl dpbo-profile add** to add an xDSL DPBO profile.

v. Run the **xdsl rfi-profile quickadd** command to quickly add an xDSL RFI profile, or run the interactive command **xdsl rfi-profile add** to add an xDSL RFI profile.

When spectrum-related profiles (except mode specific PSD profiles) are successfully configured, they can be used for activating ADSL2+ and VDSL2 ports. Mode specific PSD profiles are not directly used for activating ports but are used in spectrum-related profiles.

c. Configure service quality-related profiles.

i. Run the **xdsl inp-delay-profile quickadd** command to quickly add an xDSL INP-delay profile, or run the interactive command **xdsl inp-delay-profile add** to add an xDSL INP-delay profile.

ii. Run the **xdsl noise-margin-profile quickadd** command to quickly add an xDSL SNR margin profile, or run the interactive command **xdsl noise-margin-profile add** to add an xDSL SNR margin profile.

iii. Run the **xdsl sos-profile quickadd** command to quickly add an xDSL SOS profile, or run the interactive command **xdsl sos-profile add** to add an xDSL SOS profile.

iv. Run the **xdsl virtual-noise-profile quickadd** command to quickly add an xDSL virtual noise profile, or run the interactive command **xdsl virtual-noise-profile add** to add an xDSL virtual noise profile.

v. Run the **xdsl inm-profile quickadd** command to quickly add an xDSL impulse noise monitor profile or run the interactive command **xdsl inm-profile add** to add an xDSL pulse noise monitor profile.

NOTE

Users can determine the INP value based on the obtained INMAINPEQi and INMAIATi histogram to protect the line stability.

- **INM inter arrival time offset**: indicates the INM inter-arrival time offset (INMIATO). It determines the INMAIATi histogram parameter range with INMIATS. It also determines the start point of IAT.
- **INM inter arrival time step**: indicates the INM inter-arrival time step (INMIATS). It determines the INMAIATi histogram parameter range with INMIATO. It also determines the precision of IAT.

- **INM cluster continuation value**: indicates the INM cluster continuation (INMCC) value. It identifies a cluster and indicates the maximum number of consecutive undamaged DMT symbols allowed in a cluster.
- **INM equivalent INP mode**: Indicates the INM equivalent impulse noise protection (INP) mode. The method of calculating the equivalent INP varies according to the mode. Mode 3 is recommended because the algorithm for the mode is better than the algorithms for modes 0, 1, and 2.

After service quality-related profiles are successfully configured, they can be used for activating ADSL2+ and VDSL2 ports.

----End

Example

NOTE

The following command output is only an example. During actual configuration, the actual command output prevails.

Assume that:

- VDSL2 mode: TR129
- VDSL2 access distance: 290 m
- Profile to be configured: VDSL2 line parameter profile

Refer to the configuration described in Table 6-13. Since the access distance is smaller than 300 m, the detailed configuration procedure is as follows.

```
huawei(config)#vdsl line-profile add
{ <cr>|profile-index<U><2,770> }:6

  Command:
        vdsl line-profile add 6
  Start adding profile
  Press 'Q' to quit the current configuration and new configuration will be
neglected
>  Do you want to name the profile? (y/n) [n]:
>    Transmission mode:
>      0: Custom
>      1: All (G.992.1~5,T1.413,G.993.2)
>      2: Full rate (G.992.1/3/5,T1.413,G.993.2)
>      3: G.DMT (G.992.1/3/5,G.993.2)
>      4: G.HS (G.992.1~5,G.993.2)
>      5: ADSL (G.992.1~5,T1.413)
>      6: VDSL (G.993.2)
>    Please select (0~6) [1]:
>  Bit swap downstream 1-disable 2-enable (1~2) [2]:
>  Bit swap upstream  1-disable 2-enable (1~2) [2]:
>  Please select the form of transmit rate adaptation downstream:
>  1-fixed, 2-adaptAtStartup, 3-adaptAtRuntime, 4-adaptAtRuntimewithsos (1~4) [
2]:
>  Please select the form of transmit rate adaptation upstream:
>  1-fixed, 2-adaptAtStartup, 3-adaptAtRuntime, 4-adaptAtRuntimewithsos (1~4) [
2]:
>  Will you set SNR margin parameters? (y/n) [n]:y
>    Target SNR margin downstream (0~310 0.1dB) [60]:80     //Note that the parameter
value is expressed in 0.1 dB.
>    Minimum SNR margin downstream (0~80 0.1dB) [0]:
```

```
>   Maximum SNR margin downstream (80~310 0.1dB) [300]:
>   Target SNR margin upstream (0~310 0.1dB) [60]:80     //Note that the parameter value
is expressed in 0.1 dB.
>   Minimum SNR margin upstream (0~80 0.1dB) [0]:
>   Maximum SNR margin upstream (80~310 0.1dB) [300]:
>  Will you set DPBO parameters? (y/n) [n]:
>  Will you set UPBO parameters? (y/n) [n]:
>  Will you set power management parameters? (y/n) [n]:
>  Will you set RFI notch configuration parameter? (y/n) [n]:
>  Will you set ADSL tone blackout configuration parameter? (y/n) [n]:
>  Will you set VDSL tone blackout configuration parameter? (y/n) [n]:
>  Will you set mode-specific parameters? (y/n) [n]:y
>   Current configured modes:
>   1-defmode
>   Please select 1-Add  2-Modify  3-Save and quit   [3]:2
>   1-defmode
>   Please select  [1]:
>   G.993.2 profile:
>   1-Profile8a   2-Profile8b   3-Profile8c   4-Profile8d
>   5-Profile12a  6-Profile12b  7-Profile17a  8-Profile30a
>   Please select (1~8) [5]:7
>   VDSL2 PSD class mask:
>   1-AnnexA998-D-32                2-AnnexA998-D-64
>   3-AnnexBHPE17-M1-NUS0(B7-7)       4-AnnexB997E17-M2x-A(B7-9)
>   5-AnnexB998E17-M2x-NUS0(B8-8)     6-AnnexB998E17-M2x-NUS0-M(B8-9)
>   7-AnnexB998ADE17-M2x-NUS0-M(B8-10)  8-AnnexB998ADE17-M2x-B(B8-12)
>   9-AnnexB998ADE17-M2x-A(B8-11)     10-AnnexA998-D-48
>   11-AnnexA998-D-128              12-AnnexB998ADE17-M2x-M(B8-17)
>   Please select (1~12) [8]:9
>   VDSL2 link use of U0 1-unused, 2-used (1~2) [1]:2     //Enable US0.
>   Maximum nominal aggregate transmit power downstream
>   (-255~145 0.1dBm) [145]:
>   Maximum nominal aggregate transmit power upstream
>   (-255~145 0.1dBm) [145]:
>   Will you set PSD mask value downstream parameter? (y/n) [n]:
>   Will you set PSD mask value upstream parameter? (y/n) [n]:
>   Will you set Upstream PSD mask selection parameter? (y/n) [n]:
>   Will you set transmitter referred virtual noise parameters? (y/n) [n]:
>   Current configured modes:
>   1-defmode
>   Please select 1-Add  2-Modify  3-Save and quit   [3]:
>  Will you set network timing reference? (y/n) [n]:
>  Will you set INM parameter? (y/n) [n]:
>  Will you set SOS downstream parameter? (y/n) [n]:
>  Will you set SOS upstream parameter? (y/n) [n]:
>  Will you set the G.998.4 retransmission function? (y/n) [n]:
>  Will you set force framer setting for inp? (y/n) [n]:
  Add profile 6 successfully

huawei(config)#vdsl channel-profile add
{ <cr>|profile-index<U><2,770> }:6

  Command:
        vdsl channel-profile add 6
  Start adding profile
```

```
  Press 'Q' to quit the current configuration and new configuration will be
neglected
>  Do you want to name the profile? (y/n) [n]:
>  Data path mode 1-ATM, 2-PTM, 3-Both (1~3) [3]:
>  Will you set the minimum impulse noise protection? (y/n) [n]:y
>    Minimum impulse noise protection downstream:
>    1-noProtection    2-halfSymbol     3-singleSymbol    4-twoSymbols
>    5-threeSymbols    6-fourSymbols    7-fiveSymbols     8-sixSymbols
>    9-sevenSymbols    10-eightSymbols   11-nineSymbols    12-tenSymbols
>    13-elevenSymbols  14-twelveSymbols  15-thirteenSymbols 16-fourteenSymbols
>    17-fifteenSymbols 18-sixteenSymbols
>    Please select (1~18) [1]:4
>    Minimum impulse noise protection upstream:
>    1-noProtection    2-halfSymbol     3-singleSymbol    4-twoSymbols
>    5-threeSymbols    6-fourSymbols    7-fiveSymbols     8-sixSymbols
>    9-sevenSymbols    10-eightSymbols   11-nineSymbols    12-tenSymbols
>    13-elevenSymbols  14-twelveSymbols  15-thirteenSymbols 16-fourteenSymbols
>    17-fifteenSymbols 18-sixteenSymbols
>    Please select (1~18) [1]:4
>  Will you set interleaving delay parameters? (y/n) [n]:
>  Will you set parameters for rate? (y/n) [n]:y
>    Minimum transmit rate downstream (32~200000 Kbps) [32]:
>    Minimum reserved transmit rate downstream (32~200000 Kbps) [32]:
>    Maximum transmit rate downstream (32~200000 Kbps) [200000]:50000
>    Minimum transmit rate upstream (32~200000 Kbps) [32]:
>    Minimum reserved transmit rate upstream (32~200000 Kbps) [32]:
>    Maximum transmit rate upstream (32~200000 Kbps) [200000]:15000
>  Will you set rate thresholds? (y/n) [n]:
>  Will you set PHY-R function? (y/n) [n]:
>  Will you set erasure decoding? (y/n) [n]:
>  Will you set SOS bit rate? (y/n) [n]:
>  Will you set the G.998.4 retransmission function? (y/n) [n]:
>  Will you set channel initialization policy selection? (y/n) [n]:
  Add profile 6 successfully

huawei(config)#vdsl line-template add
{ <cr>|template-index<U><2,770> }:6

  Command:
        vdsl line-template add 6
  Start adding template
  Press 'Q' to quit the current configuration and new configuration will be
neglected
>  Do you want to name the template? (y/n) [n]:y
>    Please input template name:VDSL2-PORT1
>  Please set the line-profile index (1~770) [1]:6
>  Will you set channel configuration parameters? (y/n) [n]:y
>    Please set the channel number (1~2) [1]:1       //Configurations are required only
for channel 1.
>    Channel1 configuration parameters:
>    Please set the channel-profile index (1~770) [1]:6
  Add template 6 successfully
```

Assume that:

- VDSL2 mode: TI
- VDSL2 access distance: 290 m
- Profile to be configured: VDSL2 line parameter profile

Refer to the configuration described in Table 6-13. Since the access distance is smaller than 300 m, the detailed configuration procedure is as follows.

```
huawei(config)#vdsl service-profile add
{ <cr>|profile-index<U><2,128> }:2

  Command:
         vdsl service-profile add 2
  Start adding profile
  Press 'Q' to quit the current configuration and new configuration will be
neglected
>  Do you want to name the profile? (y/n) [n]:y
>    Please input profile name:VDSL2-PORT1
>  Data path mode 1-ATM, 2-PTM (1~2) [2]:
>  Bit swap downstream 1-enable 2-disable (1~2) [1]:
>  Bit swap upstream 1-enable 2-disable (1~2) [1]:
>  Form of transmit rate adaptation:
>    1-manual, 2-adaptAtInit, 3-dynamic (1~3) [2]:
>  Will you set parameters for rate of bearer 1? (y/n) [n]:y
>    Minimum data rate downstream (32~200000 Kbps) [32]:
>    Minimum reserved data rate downstream (32~200000 Kbps) [32]:
>    Maximum data rate downstream (32~200000 Kbps) [200000]:50000
>    Minimum data rate in low power state downstream (32~50000 Kbps) [32]:
>    Minimum data rate upstream (32~200000 Kbps) [32]:
>    Minimum reserved data rate upstream (32~200000 Kbps) [32]:
>    Maximum data rate upstream (32~200000 Kbps) [200000]:15000
>    Minimum data rate in low power state upstream (32~15000 Kbps) [32]:
>  Will you set the G.998.4 retransmission function? (y/n) [n]:
>  Will you enable bearer 2? (y/n) [n]:
  Add profile 2 successfully

huawei(config)#vdsl spectrum-profile add
{ <cr>|profile-index<U><2,128> }:2

  Command:
         vdsl spectrum-profile add 2
  Start adding profile
  Press 'Q' to quit the current configuration and new configuration will be
neglected
>  Do you want to name the profile? (y/n) [n]:y
>    Please input profile name:VDSL2-PORT1
>    Transmission mode:
>      0: Custom
>      1: All (G.992.1~5,T1.413,G.993.2)
>      2: Full rate (G.992.1/3/5,T1.413,G.993.2)
>      3: G.DMT (G.992.1/3/5,G.993.2)
>      4: G.HS (G.992.1~5,G.993.2)
>      5: ADSL (G.992.1~5,T1.413)
>      6: VDSL (G.993.2)
>    Please select (0~6) [1]:
>  Will you set ADSL tone blackout configuration parameter? (y/n) [n]:
```

```
>  Will you set VDSL tone blackout configuration parameter? (y/n) [n]:
>  Will you set RFI notch configuration parameter? (y/n) [n]:
>  Will you set the G.998.4 retransmission function? (y/n) [n]:
>  Will you set mode-specific parameters? (y/n) [n]:y
>    Current configured modes:
>    1-defmode
>    Please select 1-Add  2-Modify  3-Save and quit   [3]:2
>    1-defmode
>    Please select  [1]:
>    Will you set power management parameters? (y/n) [n]:
>    Maximum nominal aggregate transmit power downstream
>    (-255~205 0.1dBm) [200]:
>    Maximum nominal aggregate transmit power upstream
>    (-255~205 0.1dBm) [125]:
>    Maximum aggregate receive power upstream
>    value from -255(code as 0) to 255(code as 510)in steps of 1
>    (0~510 0.1dBm) [380]:
>    Will you set PSD mask value downstream parameter? (y/n) [n]:
>    Will you set PSD mask value upstream parameter? (y/n) [n]:
>    Will you set G.993.2 mode parameters? (y/n) [n]:y
>    Current configured G.993.2 modes:
>    7-17a<1>                        //The default mode for VDSL2 profiles is 17a and
therefore no change is required.
>    Please select 1-Add  2-Modify  3-Save and quit   [3]:
>    Current configured modes:
>    1-defmode
>    Please select 1-Add  2-Modify  3-Save and quit   [3]:
  Add profile 2 successfully

huawei(config)#vdsl delay-inp-profile add
{ <cr>|profile-index<U><2,128> }:2

  Command:
         vdsl delay-inp-profile add 2
  Start adding profile
  Press 'Q' to quit the current configuration and new configuration will be
neglected
>  Do you want to name the profile? (y/n) [n]:y
>    Please input profile name:VDSL2-PORT1
>  Force inp flag 1.force, 2.auto (1~2) [1]:
>  Enable or disable retransmission function in downstream of bearer 1:
>  1-enable, 2-disable (1~2) [2]:
>  Enable or disable retransmission function in upstream of bearer 1:
>  1-enable, 2-disable (1~2) [2]:
>  Will you set interleaving delay parameters of bearer 1? (y/n) [n]:
>  Will you set the minimum impulse noise protection of bearer 1? (y/n) [n]:y
//Minimum INP needs to be set only for channel 1.
>    Minimum impulse noise protection downstream:
>    1-noProtection    2-halfSymbol     3-singleSymbol    4-twoSymbols
>    5-threeSymbols    6-fourSymbols    7-fiveSymbols     8-sixSymbols
>    9-sevenSymbols    10-eightSymbols   11-nineSymbols    12-tenSymbols
>    13-elevenSymbols  14-twelveSymbols  15-thirteenSymbols 16-fourteenSymbols
>    17-fifteenSymbols 18-sixteenSymbols
>    Please select (1~18) [1]:4
>    Minimum impulse noise protection upstream:
```

```
>   1-noProtection   2-halfSymbol     3-singleSymbol    4-twoSymbols
>   5-threeSymbols   6-fourSymbols    7-fiveSymbols     8-sixSymbols
>   9-sevenSymbols   10-eightSymbols  11-nineSymbols    12-tenSymbols
>   13-elevenSymbols  14-twelveSymbols  15-thirteenSymbols 16-fourteenSymbols
>   17-fifteenSymbols 18-sixteenSymbols
>   Please select (1~18) [1]:4
>  Will you set the G.998.4 retransmission function? (y/n) [n]:
>  Enable or disable retransmission function in downstream of bearer 2:
>  1-enable, 2-disable (1~2) [2]:
>  Enable or disable retransmission function in upstream of bearer 2:
>  1-enable, 2-disable (1~2) [2]:
>  Will you set interleaving delay parameters of bearer 2? (y/n) [n]:
>  Will you set the minimum impulse noise protection of bearer 2? (y/n) [n]:
  Add profile 2 successfully

huawei(config)#vdsl noise-margin-profile add
{ <cr>|profile-index<U><2,128> }:2

  Command:
         vdsl noise-margin-profile add 2
  Start adding profile
  Press 'Q' to quit the current configuration and new configuration will be
neglected
>  Do you want to name the profile? (y/n) [n]:y
>   Please input profile name:VDSL2-PORT1
>  Will you set SNR margin parameters? (y/n) [n]:y
>   Target SNR margin downstream (0~310 0.1dB) [60]:80     //Note that the parameter
value is expressed in 0.1 dB.
>   Minimum SNR margin downstream (0~80 0.1dB) [10]:
>   Maximum SNR margin downstream (80~310 0.1dB) [310]:
>   Target SNR margin upstream (0~310 0.1dB) [60]:80     //Note that the parameter value
is expressed in 0.1 dB.
>   Minimum SNR margin upstream (0~80 0.1dB) [10]:
>   Maximum SNR margin upstream (80~310 0.1dB) [310]:
>  Will you set SRA margin parameters? (y/n) [n]:
>  Will you set rate thresholds? (y/n) [n]:
  Add profile 2 successfully
```

Assume that:

- VDSL2 mode: TR165
- VDSL2 access distance: 900 m
- Profile to be configured: VDSL2 line parameter profile

Refer to the configuration described in Table 6-13. Since the access distance is greater than 800 m, the detailed configuration procedure is as follows.

```
huawei(config)#xdsl data-rate-profile add
{ <cr>|profile-index<U><3,4294967294> }:7

  Command:
         xdsl data-rate-profile add 7
  Start adding profile
  Press 'Q' to quit the current configuration and new configuration will be
neglected
```

```
>  Do you want to set the description of the profile? (y/n) [n]:y
>    Please input profile description:VDSL2-PORT1-DS            //The limited upstream
and downstream rates are different. This profile is for limiting the downstream rate.
>  Minimum data rate (32~200000 Kbps) [32]:
>  Minimum reserved data rate (32~200000 Kbps) [32]:
>  Maximum data rate (32~200000 Kbps) [200000]:20000           ///The downstream rate
is limited to 20 Mbit/s.
>  Minimum data rate in low power state (32~20000 Kbps) [32]:
>  The ratio between L2 minimum rate and L0 rate (0~99 %) [0]:
>  Maximum data rate in low power state (32~200000 Kbps) [4000]:
>  Maximum bit error ratio 1-eminus3, 2-eminus5, 3-eminus7 (1~3) [2]:
>  Data rate threshold upshift (0~200000 Kbps) [0]:
>  Data rate threshold downshift (0~200000 Kbps) [0]:
>  Data path mode 1-ATM, 2-PTM, 3-Both (1~3) [3]:           //The default value is
recommended.
>  Will you set the G.998.4 retransmission function? (y/n) [n]:
>  Minimum SOS bit rate(Kbps) (0~65535) [8]:
  Add profile 7 successfully


huawei(config)#xdsl data-rate-profile add
{ <cr>|profile-index<U><3,4294967294> }:8

  Command:
        xdsl data-rate-profile add 8
  Start adding profile
  Press 'Q' to quit the current configuration and new configuration will be
neglected
>  Do you want to set the description of the profile? (y/n) [n]:y
>    Please input profile description:VDSL2-PORT1-US            //The limited upstream
and downstream rates are different. This profile is for limiting the upstream rate.
>  Minimum data rate (32~200000 Kbps) [32]:
>  Minimum reserved data rate (32~200000 Kbps) [32]:
>  Maximum data rate (32~200000 Kbps) [200000]:2000           //The upstream rate is
limited to 2 Mbit/s.
>  Minimum data rate in low power state (32~20000 Kbps) [32]:
>  The ratio between L2 minimum rate and L0 rate (0~99 %) [0]:
>  Maximum data rate in low power state (32~200000 Kbps) [4000]:
>  Maximum bit error ratio 1-eminus3, 2-eminus5, 3-eminus7 (1~3) [2]:
>  Data rate threshold upshift (0~200000 Kbps) [0]:
>  Data rate threshold downshift (0~200000 Kbps) [0]:
>  Data path mode 1-ATM, 2-PTM, 3-Both (1~3) [3]:           //The default value is
recommended.
>  Will you set the G.998.4 retransmission function? (y/n) [n]:
>  Minimum SOS bit rate(Kbps) (0~65535) [8]:
  Add profile 8 successfully


huawei(config)#xdsl mode-specific-psd-profile add
{ <cr>|profile-index<U><2,4294967294> }:5

  Command:
        xdsl mode-specific-psd-profile add 5
  Start adding profile
  Press 'Q' to quit the current configuration and new configuration will be
```

```
neglected
>  Do you want to set the description of the profile? (y/n) [n]:
>    Maximum nominal transmit PSD downstream
>    (300~600 -0.1dBm/Hz) [400]:
>    Maximum nominal transmit PSD upstream
>    (300~600 -0.1dBm/Hz) [380]:
>    Maximum nominal aggregate transmit power downstream
>    (-255~205 0.1dBm) [200]:
>    Maximum nominal aggregate transmit power upstream
>    (-255~205 0.1dBm) [125]:
>    Maximum aggregate receive power upstream
>    value from -255(code as 0) to 255(code as 510)in steps of 1
>    (0~510 0.1dBm) [380]:
>    Will you set PSD mask value downstream parameter? (y/n) [n]:
>    Will you set PSD mask value upstream parameter? (y/n) [n]:
>    Upstream PSD mask selection(ADSL mode):
>    1-ADLU-32/EU-32      2-ADLU-36/EU-36
>    3-ADLU-40/EU-40      4-ADLU-44/EU-44
>    5-ADLU-48/EU-48      6-ADLU-52/EU-52
>    7-ADLU-56/EU-56      8-ADLU-60/EU-60
>    9-ADLU-64/EU-64
>    Please select (1~9) [1]:
>    VDSL2 PSD mask class selection:
>    1-Class 998 Annex A or Class 997-M1c Annex B or Class 998-B Annex C
>    2-Class 997-M1x Annex B or Class 998-CO Annex C
>    3-Class 997-M2x Annex B
>    4-Class 998-M1x Annex B
>    5-Class 998-M2x Annex B
>    6-Class 998ADE-M2x Annex B
>    7-Class HPE-M1 Annex B
>    Please select (1~7) [5]:5              //According to the recommended
configurations, PSD mask is B8-6(998-M2x-B), which belongs to the classmask defined
by parameter 5.
>  Will you set VDSL2 limit PSD masks? (y/n) [n]:y
>    Current LIMITMASK for each CLASSMASK you can choose:
>    Profile8a/b/c/d:
>    1: Limit1: M2x-A               2: Limit2: M2x-B
>    3: Limit3: M2x-M               4: Limit4: M2x-NUS0
>    Profile12a/12b:
>    5: Limit1: M2x-A               6: Limit2: M2x-B
>    7: Limit3: M2x-M               8: Limit4: M2x-NUS0
>    Profile17a:
>    9: Limit1: E17-M2x-NUS0         10: Limit2: E17-M2x-NUS0-M
>    11: Limit3: E17-M2x-A
>    Profile30a:
>    12: Limit1: E30-M2x-NUS0        13: Limit2: E30-M2x-NUS0-M
>  Please select (1~13) [6]:2              //According to the recommended
configurations, PSD mask is B8-6(998-M2x-B), which belongs to the limitmask defined
by parameter 2.
>  Will you set the use of US0 for Profile8 series limit PSD mask? (y/n) [n]:y
>  The use of US0 for Profile8 series limit PSD mask:
   Limit2: M2x-B                  1-Unused  2-Used  [1]: 2
  Add profile 5 successfully

huawei(config)#xdsl line-spectrum-profile add
```

```
{ <cr>|profile-index<U><2,4294967294> }:5

  Command:
        xdsl line-spectrum-profile add 5
  Start adding profile
  Press 'Q' to quit the current configuration and new configuration will be
neglected
>  Do you want to set the description of the profile? (y/n) [n]:
>    Transmission mode:
>      0: Custom
>      1: All (G.992.1~5, T1.413, ETSI, G.993.2)
>      2: Full rate (G.992.1/3/5, T1.413, ETSI, G.993.2)
>      3: G.DMT (G.992.1/3/5, G.993.2)
>      4: G.HS (G.992.1~5, G.993.2)
>      5: ADSL (G.992.1~5, ETSI, T1.413)
>      6: VDSL2 (G.993.2)
>      7: ADSL2 & ADSL2+ (G.992.3~5)
>    Please select (0~7) [1]:
>  Will you set power management parameters? (y/n) [n]:
>  Will you set network timing reference? (y/n) [n]:
>  Bit swap downstream 1-disable 2-enable (1~2) [1]:
>  Bit swap upstream  1-disable 2-enable (1~2) [1]:
>  Will you set ADSL tone blackout configuration parameter? (y/n) [n]:
>  Will you set VDSL tone blackout configuration parameter? (y/n) [n]:
>  Minimum overhead rate upstream  (4000~248000 bps) [4000]:
>  Minimum overhead rate downstream (4000~248000 bps) [4000]:
>  Will you set G.993.2 profiles? (y/n) [n]:y
>    Current configured profiles:
>    5-Profile12a
>    Please select 1-Delete  2-Save and quit  [2]:1
>    5-Profile12a
>    Please select  [5]:
>    Current configured profiles: -
>    Please add new profiles:
>    1-Profile8a    2-Profile8b    3-Profile8c    4-Profile8d
>    5-Profile12a   6-Profile12b   7-Profile17a   8-Profile30a
>    Please select  [1]:2               //Change it the desired profile 8b.
>    Current configured profiles:
>    2-Profile8b
>    Please select 1-Delete  2-Save and quit  [2]:
>  Will you set US0 PSD masks? (y/n) [n]:
>  Optional cyclic extension flag 1-disable, 2-enable (1~2) [1]:
>  Force framer setting for inp downstream 1-false, 2-true (1~2) [1]:
>  Force framer setting for inp upstream  1-false, 2-true (1~2) [1]:
>  Will you set mode-specific parameters? (y/n) [n]:y
>    Current configured modes:
>    1-defmode
>    Please select 1-Add  2-Modify  3-Save and quit  [3]:2
>    1-defmode
>    Please select  [1]:
>    Please select the mode specific PSD profile index (1~4294967294) [1]:5      //Use
the configured mode specific PSD profile 5.
>    Current configured modes:
>    1-defmode
>    Please select 1-Add  2-Modify  3-Save and quit  [3]:
```

```
>  Will you set the G.998.4 retransmission function? (y/n) [n]:
  Add profile 5 successfully

huawei(config)#xdsl inp-delay-profile add
{ <cr>|profile-index<U><2,4294967294> }:2

  Command:
         xdsl inp-delay-profile add 2
  Start adding profile
  Press 'Q' to quit the current configuration and new configuration will be
neglected
>  Do you want to set the description of the profile? (y/n) [n]:
>  Will you set the minimum impulse noise protection transported over DMT
>    symbols with a subcarrier spacing of 4.3125 KHz? (y/n) [n]:y       //In 8b profile,
Tone Spacing is 4.3125 KHz. Set the minimum INP.
>    Minimum impulse noise protection downstream:
>    1-noProtection    2-halfSymbol     3-singleSymbol    4-twoSymbols
>    5-threeSymbols    6-fourSymbols    7-fiveSymbols     8-sixSymbols
>    9-sevenSymbols    10-eightSymbols   11-nineSymbols    12-tenSymbols
>    13-elevenSymbols  14-twelveSymbols  15-thirteenSymbols 16-fourteenSymbols
>    17-fifteenSymbols 18-sixteenSymbols
>    Please select (1~18) [1]:4
>    Minimum impulse noise protection upstream:
>    1-noProtection    2-halfSymbol     3-singleSymbol    4-twoSymbols
>    5-threeSymbols    6-fourSymbols    7-fiveSymbols     8-sixSymbols
>    9-sevenSymbols    10-eightSymbols   11-nineSymbols    12-tenSymbols
>    13-elevenSymbols  14-twelveSymbols  15-thirteenSymbols 16-fourteenSymbols
>    17-fifteenSymbols 18-sixteenSymbols
>    Please select (1~18) [1]:4
>  Will you set the minimum impulse noise protection transported over DMT
>    symbols with a subcarrier spacing of 8.625 KHz? (y/n) [n]:
>  Will you set interleaving delay parameters? (y/n) [n]:
>  Maximum delay variation, it ranges from 0.1 to 25.4 in steps of 0.1 ms
>  A special value 255 indicates that no delay variation bound is imposed
>  (1~255 0.1ms) [255]:
>  Channel initialization policy selection (0~2) [0]:
>  Will you set the G.998.4 retransmission function? (y/n) [n]:
  Add profile 2 successfully

huawei(config)#xdsl noise-margin-profile add
{ <cr>|profile-index<U><2,4294967294> }:2

  Command:
         xdsl noise-margin-profile add 2
  Start adding profile
  Press 'Q' to quit the current configuration and new configuration will be
neglected
>  Do you want to set the description of the profile? (y/n) [n]:
>  Will you set SNR margin parameters? (y/n) [n]:y
>    Target SNR margin downstream (0~310 0.1dB) [60]:80     //Note that the parameter
value is expressed in 0.1 dB.
>    Minimum SNR margin downstream (0~80 0.1dB) [10]:
>    Maximum SNR margin downstream (80~310 0.1dB) [310]:
>    Target SNR margin upstream (0~310 0.1dB) [60]:80     //Note that the parameter value
is expressed in 0.1 dB.
```

```
>   Minimum SNR margin upstream (0~80 0.1dB) [10]:
>   Maximum SNR margin upstream (80~310 0.1dB) [310]:
> Will you set signal-to-noise ratio mode parameters? (y/n) [n]:
> Please select the form of transmit rate adaptation downstream:
> 1-fixed, 2-adaptAtStartup, 3-adaptAtRuntime (1~3) [2]:
> Please select the form of transmit rate adaptation upstream:
> 1-fixed, 2-adaptAtStartup, 3-adaptAtRuntime (1~3) [2]:
 Add profile 2 successfully
```

Configuring VDSL2 Line Bonding

To ensure longer access distance at the same access rate or higher access rate in the same access distance, configure VDSL2 line bonding.

Prerequisites

- The port to be bound has no service flow.
- The port to be bound is in the activating or deactivated state.

NOTE
An xDSL port can be in any of the following states: activating, activated, deactivated, and loopback.

Procedure

Create a bonding group.

In global config mode, run the **bonding-group add** command to create a bonding group.

Key parameters:

- **primary-port**: indicates the primary port in the bonding group. After a bonding group is created, service flows can be created only on the primary port.
- **scheme**: indicates the local bonding mode, which can be **ATM**, **EFM**, or **TDIM**. For a VDSL2 PTM bonding group, the local bonding mode must be set to **EFM**.
- **peer-scheme**: indicates the peer bonding mode, which must be the same as **scheme**.

Step 1 Add member ports for a bonding group.

Run the **bonding-group link add** command to add a member port.

NOTE
One member port is added each time this command is executed.

Step 2 (Optional) Create a bonding group profile.

Run the **xdsl bonding-group-profile add** command to create a bonding group profile and set line parameters for ports in the bonding group.

- There is a default profile: profile 1.
- The priority of the bonding group profile is higher than the line parameter profiles of the ports in the bonding group. When both the bonding group profile and line parameter profiles of the ports are used, the bonding group profile takes effect. If the maximum and minimum upstream/downstream transmission rates are set to 0, the rates are not limited in the bonding group profile and are determined by the rate limits specified in the line parameter profiles of the ports.

Step 3 Activate a bonding group.

Run the **active bonding-group** command to activate a bonding group.

Step 4 Query information about a bonding group.

Run the **display bonding-group** command to query information about a bonding group.

----End

Example

To add VDSL2 ports 0/2/0 and 0/2/1 to bonding group 1 (0/2/0 is the primary port) and activate the bonding group using bonding group profile 1, do as follows:

```
huawei(config)#bonding-group add 1 primary-port 0/2/0 scheme efm peer-scheme efm
huawei(config)#bonding-group link add 1 0/2/1
huawei(config)#active bonding-group 1 profile-index 1
```

Configuring VDSL2 User Ports

xDSL ports must be activated before they are used to transmit services. This topic describes how to activate VDSL2 ports and enables the ports to use VDSL2 profiles.

Prerequisites

Overview of Configuring VDSL2 Templates and Profiles has been completed based on the data plan.

Procedure

- Do as follows to configure the VDSL2 user ports when the VDSL2 mode is TR129:
  a. In global config mode, run the **interface vdsl** command to enter the VDSL mode.
  b. Run the **deactivate** command to deactivate VDSL2 ports.
  c. Run the **activate** command to activate VDSL2 ports and enable them to use the VDSL2 line template.
  d. Run the **alarm-config** command to enable the VDSL2 ports to use the VDSL2 alarm template.
- Do as follows to configure the VDSL2 user ports when the VDSL2 mode is TI:
  a. In global config mode, run the **interface vdsl** command to enter the VDSL mode.
  b. Run the **deactivate** command to deactivate VDSL2 ports.
  c. Run the **activate** command to activate VDSL2 ports and enable them to use VDSL2 line parameter profiles.
  d. Run the **alarm-config** command to enable the VDSL2 ports to use the VDSL2 alarm template.
- Do as follows to configure the VDSL2 user ports when the VDSL2 mode is TR165:
  a. In global config mode, run the **interface vdsl** command to enter the VDSL mode.
  b. Run the **deactivate** command to deactivate VDSL2 ports.
  c. Run the **activate** command to activate VDSL2 ports and enable them to use VDSL2 line parameter profiles.
  d. Run the **alarm-config** command to enable the VDSL2 ports to use the VDSL2 alarm template.

----End

Example

In TR129 mode, to activate VDSL2 port 0/2/0 and enable the port to use VDSL2 alarm template 3 configured in the "Example" section of Configuring a VDSL2 Alarm Template and VDSL2 line template 6 configured in the "Example" section of Configuring a VDSL2 Line Parameter Profile, do as follows:

```
huawei(config)#interface vdsl 0/2
huawei(config-if-vdsl-0/2)#deactivate 0
huawei(config-if-vdsl-0/2)#activate 0 template-index 6
huawei(config-if-vdsl-0/2)#alarm-config 0 3
```

Configuring A/V Adaptation for VDSL2 Lines

VDSL2 is compatible with ADSL, ADSL2, and ADSL2+. Hence, in addition to the VDSL2 mode, a VDSL2 line can be activated in ADSL, ADSL2, or ADSL2+ mode. The activation mode can be set to A/V adaptation for a VDSL2 line so that the line can adapt to a proper activation mode according to the type of the connected modem.

Context

The A/V adaptation process of a VDSL2 line is as follows:

1. During the activation of a VDSL2 port, the training is initiated on the central office (CO) device and customer premises equipment (CPE). During the training, CO and CPE devices exchanges their capability information (that is, the 6.3.2 Annex Types and US/DS Frequency Band Planning). Different transmission modes have different priorities. For example, the priority of VDSL2 (G.993.2) is higher than that of ADSL2+ (G.992.5). Based on the intersection capabilities, CO and CPE devices select an optimal transmission mode for negotiation and then line activation after a successful negotiation. If the line can be activated, the negotiation stops. Otherwise, CO and CPE devices select the transmission mode with the next priority level for negotiation. This process repeats until the negotiation succeeds and the port is activated. Hence, to achieve A/V adaptation, ensure that the **Transmission mode** specified for the CO device includes all VDSL2, ADSL, ADSL2, and ADSL2+ standards and Annex types.
2. ADSL/ADSL2/ADSL2+ and VDSL2 use different packet encapsulation modes. Therefore, the packet encapsulation mode must be configured using either of the following methods:
   - Traditional configuration: The ports activated in VDSL2 mode are encapsulated in PTM mode and those activated in ADSL/ADSL2/ADSL2+ mode are encapsulated in ATM mode.

     In traditional configuration, to implement A/V adaptation, one PTM service flow and one ATM service flow must be configured for one VDSL2 port. After the configuration, the MA5600T/MA5603T/MA5608T determines which service flow takes effect based on the port activation mode, ensuring successful user service access.
   - Special configuration: The ports activated in VDSL2 or ADSL2/ADSL2+ mode are encapsulated in PTM mode. This configuration applies when the DSLAM matches the upper-layer device or OSS for special service connections. This configuration cannot be used if the ports are activated in ADSL mode. In this case, use the traditional configuration.

In special configuration, only one PTM service flow needs to be configured for one VDSL2 port.

The following describes how to set the transmission mode and configure PTM and ATM service flows.

NOTE

In addition to enabling a line to adapt to a proper activation mode according to the type of the connected modem, the A/V adaptation function can also be used in a long-distance VDSL2 transmission scenario. In this scenario, even if VDSL2 modems are used as terminals, the line may fail to be activated in VDSL2 mode because of poor line quality. If A/V adaptation is not enabled, line activation fails; if A/V adaptation is enabled and the VDSL2 modems can work in ATM mode, the line can be activated in ADSL, ADSL2, or ADSL2+ mode. This scenario is rare and is not recommended. If the transmission distance is longer than 1.2 km, ADSL2+ access mode is recommended. In the long-distance VDSL2 transmission scenario, A/V adaptation must be configured at the CO device and corresponding configuration must be made on VDSL2 modems. That is, PTM and ATM service flows must be configured. For details about the configuration procedures, see the user guide for the modem.

Procedure

When configuring the line parameter profiles, ensure that the **Transmission mode** includes all VDSL2, ADSL, ADSL2, and ADSL2+ standards and Annex types, for example, the default value "1: All (G.992.1~5,T1.413,G.993.2)".

In TR129 mode, run the **vdsl line-profile add** command to set **Transmission mode**; in TI mode, run the **vdsl spectrum-profile add** command to set **Transmission mode**; in TR165 mode, run the **xdsl line-spectrum-profile add** command to set **Transmission mode**.

Step 1 If traditional configuration is used, run the **service-port** command to configure one PTM service flow and one ATM service flow for one VDSL2 port. If special configuration is used, in diagnosis mode, run the **xdsl adsl-ptm-mode enable** command to enable ADSL PTM globally. Then, in global config mode, run the **service-port** command to configure one PTM service flow for one VDSL2 port.

----End

Example

NOTE

The following command output is only an example. During actual configuration, the actual command output prevails.

To configure A/V adaptation for VDSL2 port 0/2/0 in TR129 mode (the line parameters are the same as those in the "Example" section of Configuring a VDSL2 Line Parameter Profile), do as follows:

```
//Configuring VDSL2 line parameter profile
huawei(config)#vdsl line-profile add
{ <cr>|profile-index<U><2,770> }:6

  Command:
         vdsl line-profile add 6
  Start adding profile
  Press 'Q' to quit the current configuration and new configuration will be
neglected
>  Do you want to name the profile? (y/n) [n]:
>    Transmission mode:
>      0: Custom
```

```
>     1: All (G.992.1~5,T1.413,G.993.2)
>     2: Full rate (G.992.1/3/5,T1.413,G.993.2)
>     3: G.DMT (G.992.1/3/5,G.993.2)
>     4: G.HS (G.992.1~5,G.993.2)
>     5: ADSL (G.992.1~5,T1.413)
>     6: VDSL (G.993.2)
>   Please select (0~6) [1]:         //Select the default value, which includes all
VDSL2, ADSL, ADSL2, and ADSL2+ standards and Annex types.
>  Bit swap downstream 1-disable 2-enable (1~2) [2]:
>  Bit swap upstream  1-disable 2-enable (1~2) [2]:
>  Please select the form of transmit rate adaptation downstream:
>  1-fixed, 2-adaptAtStartup, 3-adaptAtRuntime, 4-adaptAtRuntimewithsos (1~4) [
2]:
>  Please select the form of transmit rate adaptation upstream:
>  1-fixed, 2-adaptAtStartup, 3-adaptAtRuntime, 4-adaptAtRuntimewithsos (1~4) [
2]:
>  Will you set SNR margin parameters? (y/n) [n]:y
>   Target SNR margin downstream (0~310 0.1dB) [60]:80     //Note that the parameter
value is expressed in 0.1 dB.
>   Minimum SNR margin downstream (0~80 0.1dB) [0]:
>   Maximum SNR margin downstream (80~310 0.1dB) [300]:
>   Target SNR margin upstream (0~310 0.1dB) [60]:80     //Note that the parameter value
is expressed in 0.1 dB.
>   Minimum SNR margin upstream (0~80 0.1dB) [0]:
>   Maximum SNR margin upstream (80~310 0.1dB) [300]:
>  Will you set DPBO parameters? (y/n) [n]:
>  Will you set UPBO parameters? (y/n) [n]:
>  Will you set power management parameters? (y/n) [n]:
>  Will you set RFI notch configuration parameter? (y/n) [n]:
>  Will you set ADSL tone blackout configuration parameter? (y/n) [n]:
>  Will you set VDSL tone blackout configuration parameter? (y/n) [n]:
>  Will you set mode-specific parameters? (y/n) [n]:y
>   Current configured modes:
>   1-defmode
>   Please select 1-Add  2-Modify  3-Save and quit   [3]:2
>   1-defmode
>   Please select  [1]:
>   G.993.2 profile:
>   1-Profile8a   2-Profile8b   3-Profile8c   4-Profile8d
>   5-Profile12a  6-Profile12b  7-Profile17a  8-Profile30a
>   Please select (1~8) [5]:7
>   VDSL2 PSD class mask:
>   1-AnnexA998-D-32                2-AnnexA998-D-64
>   3-AnnexBHPE17-M1-NUS0(B7-7)       4-AnnexB997E17-M2x-A(B7-9)
>   5-AnnexB998E17-M2x-NUS0(B8-8)     6-AnnexB998E17-M2x-NUS0-M(B8-9)
>   7-AnnexB998ADE17-M2x-NUS0-M(B8-10)  8-AnnexB998ADE17-M2x-B(B8-12)
>   9-AnnexB998ADE17-M2x-A(B8-11)     10-AnnexA998-D-48
>   11-AnnexA998-D-128              12-AnnexB998ADE17-M2x-M(B8-17)
>   Please select (1~12) [8]:9
>   VDSL2 link use of U0 1-unused, 2-used (1~2) [1]:2     //Enable US0.
>   Maximum nominal aggregate transmit power downstream
>   (-255~145 0.1dBm) [145]:
>   Maximum nominal aggregate transmit power upstream
>   (-255~145 0.1dBm) [145]:
>   Will you set PSD mask value downstream parameter? (y/n) [n]:
```

```
>   Will you set PSD mask value upstream parameter? (y/n) [n]:
>   Will you set Upstream PSD mask selection parameter? (y/n) [n]:
>   Will you set transmitter referred virtual noise parameters? (y/n) [n]:
>   Current configured modes:
>   1-defmode
>   Please select 1-Add  2-Modify  3-Save and quit   [3]:
>  Will you set network timing reference? (y/n) [n]:
>  Will you set INM parameter? (y/n) [n]:
>  Will you set SOS downstream parameter? (y/n) [n]:
>  Will you set SOS upstream parameter? (y/n) [n]:
>  Will you set the G.998.4 retransmission function? (y/n) [n]:
>  Will you set force framer setting for inp? (y/n) [n]:
  Add profile 6 successfully
huawei(config)#vdsl channel-profile add
{ <cr>|profile-index<U><2,770> }:6

  Command:
        vdsl channel-profile add 6
  Start adding profile
  Press 'Q' to quit the current configuration and new configuration will be
neglected
>  Do you want to name the profile? (y/n) [n]:
>  Data path mode 1-ATM, 2-PTM, 3-Both (1~3) [3]:
>  Will you set the minimum impulse noise protection? (y/n) [n]:y
>   Minimum impulse noise protection downstream:
>   1-noProtection   2-halfSymbol     3-singleSymbol    4-twoSymbols
>   5-threeSymbols   6-fourSymbols    7-fiveSymbols     8-sixSymbols
>   9-sevenSymbols   10-eightSymbols   11-nineSymbols    12-tenSymbols
>   13-elevenSymbols  14-twelveSymbols  15-thirteenSymbols 16-fourteenSymbols
>   17-fifteenSymbols 18-sixteenSymbols
>   Please select (1~18) [1]:4
>   Minimum impulse noise protection upstream:
>   1-noProtection   2-halfSymbol     3-singleSymbol    4-twoSymbols
>   5-threeSymbols   6-fourSymbols    7-fiveSymbols     8-sixSymbols
>   9-sevenSymbols   10-eightSymbols   11-nineSymbols    12-tenSymbols
>   13-elevenSymbols  14-twelveSymbols  15-thirteenSymbols 16-fourteenSymbols
>   17-fifteenSymbols 18-sixteenSymbols
>   Please select (1~18) [1]:4
>  Will you set interleaving delay parameters? (y/n) [n]:
>  Will you set parameters for rate? (y/n) [n]:y
>   Minimum transmit rate downstream (32~200000 Kbps) [32]:
>   Minimum reserved transmit rate downstream (32~200000 Kbps) [32]:
>   Maximum transmit rate downstream (32~200000 Kbps) [200000]:50000
>   Minimum transmit rate upstream (32~200000 Kbps) [32]:
>   Minimum reserved transmit rate upstream (32~200000 Kbps) [32]:
>   Maximum transmit rate upstream (32~200000 Kbps) [200000]:15000
>  Will you set rate thresholds? (y/n) [n]:
>  Will you set PHY-R function? (y/n) [n]:
>  Will you set erasure decoding? (y/n) [n]:
>  Will you set SOS bit rate? (y/n) [n]:
>  Will you set the G.998.4 retransmission function? (y/n) [n]:
>  Will you set channel initialization policy selection? (y/n) [n]:
  Add profile 6 successfully
huawei(config)#vdsl line-template add
{ <cr>|template-index<U><2,770> }:6
```

```
 Command:
       vdsl line-template add 6
 Start adding template
 Press 'Q' to quit the current configuration and new configuration will be
neglected
>  Do you want to name the template? (y/n) [n]:y
>   Please input template name:VDSL2-PORT1
>  Please set the line-profile index (1~770) [1]:6
>  Will you set channel configuration parameters? (y/n) [n]:y
>   Please set the channel number (1~2) [1]:1      //Configurations are required only
for channel 1.
>   Channel1 configuration parameters:
>   Please set the channel-profile index (1~770) [1]:6
 Add template 6 successfully

//Configuring VDSL2 user port
huawei(config)#interface vdsl 0/2
huawei(config-if-vdsl-0/2)#deactivate 0
huawei(config-if-vdsl-0/2)#activate 0 template-index 6
huawei(config-if-vdsl-0/2)#alarm-config 0 1    //Use the default alarm template 1.
huawei(config-if-vdsl-0/2)#quit

//Configuring service ports (Run the following commands in traditional configuration.)
huawei(config)#service-port 2 vlan 100 vdsl mode ptm 0/2/0   //Configure a PTM service
flow.
huawei(config)#service-port 3 vlan 100 vdsl mode atm 0/2/0 vpi 0 vci 35   //Configure
an ATM service flow.

//Configuring service ports (Run the following command in special configuration.)
huawei(config)#diagnose
huawei(diagnose)%%xdsl adsl-ptm-mode enable
huawei(diagnose)%%config
huawei(config)#service-port 2 vlan 100 vdsl mode ptm 0/2/0   //Configure a PTM service
flow.
```

6.5.5 VDSL2 Maintenance and Fault Diagnosis

There are many maintenance and fault diagnosis methods for DSL lines. The following describes the common faults and troubleshooting methods.

Common VDSL2 Line Faults and Troubleshooting Methods

The diagnosis and troubleshooting methods for common VDSL2 line faults are described to facilitate line maintenance.

Common Faults on VDSL2 Lines

1. When the line is activated for the first time,

- The line fails to be activated.
- The activation rate is slow.

2. When the line is normal operation, the line quality degrades and consequently the line rate decreases or even the line is deactivated.

Alarms and events involved in these faults are as follows:

- 0x29100001 The ring topology in the subscriber port is found
- 0x3d300003 The VDSL port is automatically deactivated due to loss of signal(LOS) or loss of frame(LOF)
- 0x3d300006 The line performance statistics of the VDSL port reach the threshold
- 0x3d300007 The xDSL channel downstream rate is lower than the threshold
- 0x3d300009 It fails to activate the port by using the VDSL line configuration parameters
- 0x3d30000a The channel performance statistics of the VDSL port reach the threshold
- 0x3d30000b The xDSL channel upstream rate is lower than the threshold
- 0x3d30001a The VDSL port activated rate change

Causes of the Common Faults

Table 6-14 Causes of the common VDSL2 line faults

Reason	Description	Troubleshooting
Physical lines are of poor quality.	There are engineering issues. For example, the physical line is not securely connected or deteriorates.	1. Resolve the engineering issues by referring to 6.5.2 VDSL2 Engineering Precautions. 2. In global config mode, run the **display event history** command to check if the related events have been generated. If yes, clear the event by referring to the Alarm and Event Handling.
	There is a loop in subscriber lines.	In global config mode, run the **display alarm history alarmid 0x29110001** command to check if a loop alarm has been generated. If yes, communicate with the subscriber that owns the alarming port and help the subscriber check its line connections and release the loop.
	There are interference sources around DSL lines.	Check if there are strong interference sources around subscriber lines, such as a wireless base station and high-frequency switch-mode power supply. 1. Remove the interference sources as much as possible or reroute the subscriber lines. 2. You can also deal with the interference by RFI Notching, Tone Blackout, increasing SNR margin, or limiting the activation rate.
	The VDSL2 board or port is faulty.	Rectify the fault by referring to Loopback on a VDSL2 Port.
The modem malfunctions.	The performance of the modem is poor or the modem is unstable, or the modem is faulty.	First, reset the modem; if noneffective, replace the modem.
Line parameters are improperly	US0 is not enabled for a long line.	Enable US0 for a long line (such as a line with a length more than 500 m). 1. For the TR129 and TI modes, run the **display parameter** command in the VDSL mode to check if the value of **VDSL2 link use of U0** is Used; for

Reason	Description	Troubleshooting
configured.		the TR165 mode, run the **display xdsl mode-specific-psd-profile** *profile-index* command to check if the value of **US0 config for VDSL2 PSD LIMITMASK** is Used. If not, enable US0 by referring to Configuring a VDSL2 Line Parameter Profile and then reactivate the port using the new profile. 2. In global config mode, run the **display event history** command to check if the related events have been generated. If yes, clear the event by referring to the Alarm and Event Handling.
	The target SNR margin is improperly configured. A large margin may decrease the activation rate and a small margin may affect the stability of the line.	1. In VDSL mode, run the **display line operation** command to check if the value of **Line SNR margin downstream/upstream** is proper compared with the historical values or the value of a functional port. If the value is improper, follow instructions provided in Configuring a VDSL2 Line Parameter Profile to modify SNR Margin configurations. Then reactivate the port using the new profile. 2. In global config mode, run the **display event history** command to check if the related events have been generated. If yes, clear the event by referring to the Alarm and Event Handling.
	The minimum INP is improperly configured. There is a restrictive relationship between INP and line activation rate. Under a certain interleave depth, the line activation rate decreases with the increase of the INP value. If the minimum INP is large (for example, 16), the maximum interleave delay must also be large (for example, 63 ms). If the minimum INP is large while the maximum interleave delay is small, the line activation rate will be low or even the activation fails.	1. In VDSL mode, run the **display parameter** command to check if the values of **Minimum impulse noise protection downstream/upstream** and **Maximum interleaving delay downstream/upstream** are proper. If the values are improper, follow the instructions provided in Configuring a VDSL2 Line Parameter Profile to modify the configurations of the minimum INP and maximum interleave delay. Then reactivate the port using the new profile. 2. In global config mode, run the **display event history** command to check if the related events have been generated. If yes, clear the event by referring to the Alarm and Event Handling.

Loopback on a VDSL2 Port

This section describes how to perform a loopback on a very-high-speed digital subscriber line 2 (VDSL2) port to locate a VDSL2 service fault. A loopback on a VDSL2 port can be performed to determine whether the service board housing the VDSL2 port is communicating with the backplane properly.

Prerequisites

- The VDSL2 port is deactivated.
- The VDSL2 service ran properly before the fault occurred. (This confirms that a downstream service flow exists between the control board and the VDSL2 service board).

Impact on the System

- When a VDSL2 port is executing loopback operations, the port cannot forward packets properly, and all services carried on the port are interrupted.
- If a VDSL2 port is not isolated before executing loopback operations, a broadcast storm may occur on the device and affect services carried on other ports.

You must, therefore, set a loopback duration before starting the loopback, or run the **undo loopback** command to cancel the loopback immediately after it is complete.

Procedure

Run the **loopback** command in VDSL mode to start a loopback on a VDSL2 port.

NOTE
Port loopback is classified as local loopback and remote loopback. For details about local loopback and remote loopback, see section *Reference* in the following section. VDSL2 ports support only local loopback.

For example, run the following command to start a local loopback on port 0/1/0:

```
huawei(config-if-vdsl-0/1)#loopback 0 local
```

Step 1 If the VDSL board is working in asynchronous transfer mode (ATM), run the **atm-ping** command in VDSL mode to check the connectivity of the loopback channel. If the VDSL board is working in packet transfer mode (PTM), use an external testing device, such as the SmartBits, to check the connectivity of the loopback channel by sending packets to the service board.

If, for example, the virtual path identifier (VPI) and virtual channel identifier (VCI) of the tested service flow on port 0/1/0 is 0/35, and the port is working in ATM mode, run the following command to check the connectivity of the loopback channel set up in Step 1:

```
huawei(config-if-vdsl-0/1)#atm-ping 0 0 35
```

NOTE
- If the ping operation is successful and no packets are lost, the loopback channel is connected.
- If the ping operation fails, the channel is broken.
- If the ping operation is successful but some packets are lost, the channel is faulty.

Step 2 Run the **undo loopback** command to cancel the loopback after the loopback operation is complete.

NOTE
A port on which a loopback is being performed cannot be activated.

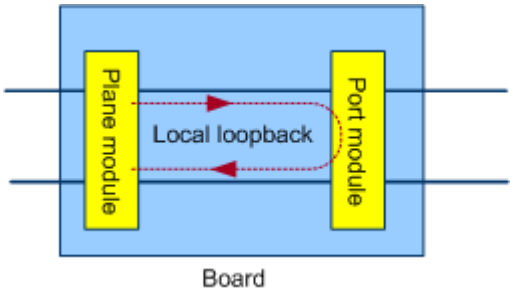
----End

Reference

Introduction to local loopback

Local loopback, also called inloop, near-end loopback, or central office (CO) loopback, is a loopback performed from the port processing module of a service board to the backplane. In this loopback, signals are sent from the backplane to the port processing module, and then be sent back to the backplane. The following figure shows a local loopback.

Figure 6-35 Local loopback

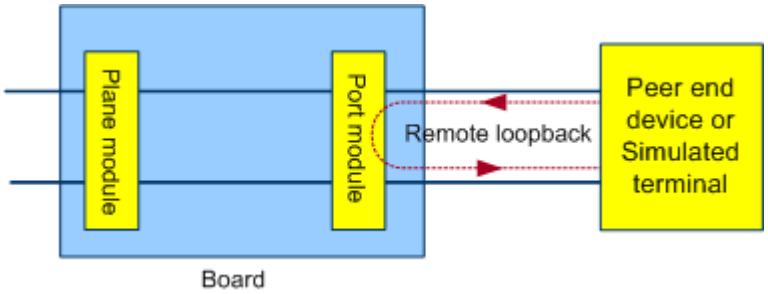


A local loopback checks whether the service channel between the control board and the service board is working properly. When a service failure occurs, this operation can be used to locate faults that occur on the control board or on the logic chip or board chipset of a service board.

Remote Loopback

Remote loopback, also called outloop, refers to the loopback from the port processing module inside the board to the subscriber line. In remote loopback, the signals between the user-side device (such as the modem) and the port signal receiving module directly return to the user-side device through the port signal sending module over the subscriber line. The test aims to check whether the upstream service between the customer premises equipment (CPE) and the board is through, and whether packet loss exists. When the service failure occurs, the fault is located on the CPE or the board chip set. The following figure shows the remote loopback.

Figure 6-36 Remote loopback



6.6 VDSL2 Reference Standards and Protocols

The reference standards and protocols of the VDSL2 feature are as follows:

Table 6-15 Reference standards and protocols of the VDSL2 feature

Standard No.	Description
ITU-T G.993.2	Very-high-speed digital subscriber line transceivers 2 (VDSL2)
ITU-T G.997.1	Physical layer management for digital subscriber line (DSL) transceivers
ITU-T G.998.2	Ethernet-based multi-pair bonding
ITU-T G.998.4	Improved Impulse Noise Protection (INP) for DSL Transceivers
ITU-T G.994.1	Handshake procedures for digital subscriber line (DSL) transceivers
Broadband Forum TR-129	Protocol-Independent Management Model for Next Generation DSL Technologies
Broadband Forum TR-165	Vector of Profiles
Broadband Forum TR-159	Management Framework for xDSL Bonding

6.7 Appendix 1: Introduction to the VDSL2 Coding/Decoding Technologies

VDSL2 coding/decoding is essential for improving line quality and performance.

DMT Modulation

DMT divides transmission bandwidth into n stand-alone or discrete sub-carriers (also called tones) and performs orthogonal transforming on data segments in each sub-carrier. The most common transforming method is discrete Fourier transform (DFT). The data rate of each sub-carrier is 1/n of the entire data rate.

Pilot Tone

DMT requires strict clock synchronization between devices at both ends. For clock synchronization, several pilot tones can be inserted to avoid wandering of frequency points.

Optional Cyclic Extension Length

DMT supports a cyclic extension between DMT symbols and uses the cyclic extension for protection. This cyclic extension is also called cyclic prefix. A cyclic prefix eliminates the

interference caused by latency extension between DMT symbols but lowers the bandwidth usage.

ITU-T Recommendation G.993.2 stipulates calculation of optional cyclic extension length. Specifically, if the path conditions are unfavorable, the cyclic prefix can be extended to prolong the protection interval, which helps eliminate interference between DMT symbols. If the path conditions are favorable, the cyclic prefix can be narrowed to increase bandwidth usage.

The Huawei access device enables users to run commands to set **Optional Cyclic Extension Flag** (enabled or disabled), which complies with ITU-T Recommendation G.997.1. **Optional Cyclic Extension Flag** identifies whether to enable the optional cyclic extension. If it is enabled, the algorithm for calculating the optional cyclic prefix is started; if it is disabled, the cyclic prefix of a fixed length is used.

Scrambling

Data transmitted over the line may contain long strings of consecutive 0s or 1s. Such data may interfere with the data of adjacent lines and cause incorrect or difficult delimitation on the peer device. The long strings of consecutive 0s or 1s must be processed to appear randomly generated before signals are carried over a line. This is the purpose of scrambling.

Scrambling generally involves inserting a fixed-length sequence at the local end and removing the sequence at the remote end. This inserted sequence keeps the signals stochastic over a line.

Trellis Coding

Common path coding techniques can be classified into convolutional coding and block coding. Trellis coding is a code modulation technique that combines convolutional coding with the digital modulation mode. The corresponding decoding technique is called Viterbi decoding.

The process of Trellis coding entails the redundancy of only one bit. Hence, Trellis coding features a higher coding efficiency and a simplified coding mechanism. However, the corresponding Viterbi decoding has a complicated process. Viterbi decoding can be divided into hard decision (HD) and soft decision (SD). SD adds some probability weighted calculation to the decoding process and thus Viterbi decoding has a stronger error correcting capability.

Trellis coding is mainly targeted at burst errors. It can correctly parse the discrete error bits in the transmission and features strong code gaining and error correcting capabilities. The VDSL2 standard defines Trellis coding as mandatory for VDSL2 implementation.

FEC

In general, there are multiple error correction mechanisms. Some depend on the transmission system itself to check the data and correct the errors after the data arrives at the peer end. Others only check the data and do not correct the errors; if any error is detected, the data is retransmitted. Forward error correction (FEC) belongs to the former category and applies to real-time services, as such services do not tolerate the latency caused by retransmission. FEC is not exclusive to DSL and is commonly used for error correction.

When applied in DSL, FEC uses Reed-Solomon (RS) coding and appends redundancy bytes to the original data. These redundancy bytes identify and correct errors. All error correction mechanisms have a trade-off in performance; accordingly, FEC sacrifices some bandwidth when implemented.

7 Vectoring

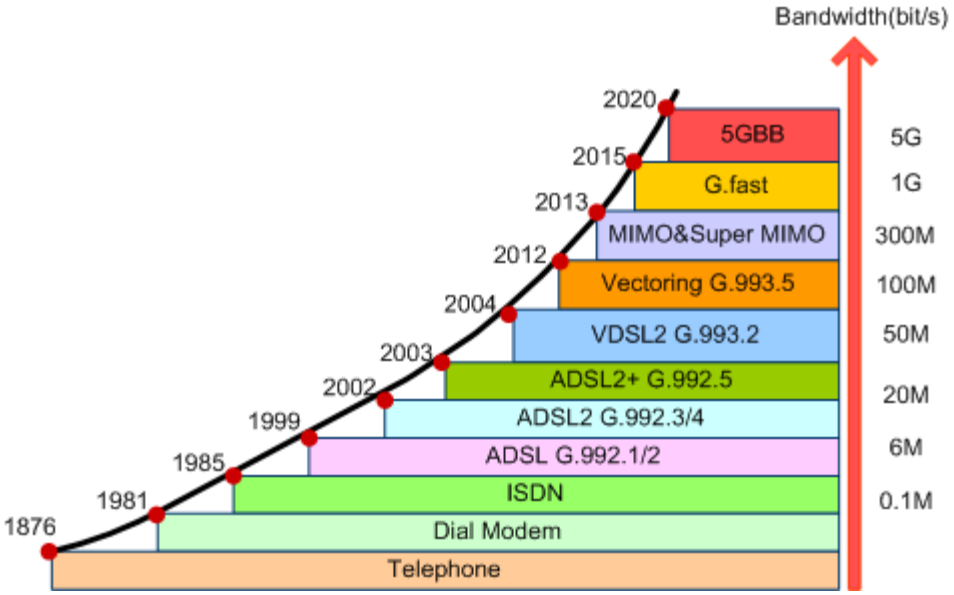
About This Chapter

Vectoring is a technology that uses vectoring algorithms to cancel crosstalk in multi-pair VDSL2 lines, thereby improving VDSL2 line bandwidths.

7.1 Background

DSL has become the most popular fixed broadband access technology since the birth of this technology in 1990s. In addition, the DSL technology was continuously seeking for technical breakthroughs and optimization, owing to which the services supported by DSL have been enriched from the original data service only to multi-play services, such as high-speed Internet (HSI), IPTV, VoIP, private line access, and mobile service bearing.

Figure 7-1 Copper line technology evolution



Key copper line technologies are as follows:

- ADSL2+: supports mainstream services represented by HSI, IPTV, and VoIP services. However, ADSL2+ cannot support high-bandwidth services, such as high definition television (HDTV), mass file sharing, and interactive multimedia. These services require an advanced DSL technology, VDSL2, for higher transmission rates.
- VDSL2: complies with ITU-T Recommendation G.993.2 and is an extension of VDSL1 that complies with ITU-T Recommendation G.993.1. VDSL2 is compatible with ADSL, ADSL2, and ADSL2+, but is not compatible with the less-common VDSL1. Primary targets of broadcast access network construction, "smooth evolution, low overall costs, fast go-to-market, and manageable and controllable", promote the development of FTTx networks, including fiber to the building (FTTB), fiber to the curb (FTTC), and fiber to the home (FTTH) networks. In FTTx networks, VDSL2 is the mainstream copper line access mode in the "last mile" because of its high bandwidth (ideally, 100 Mbit/s) in a short distance (within 1.2 km).

  VDSL2 uses high frequencies and therefore crosstalk between lines is prominently obvious. Compared with the bandwidth of single-pair VDSL2 access, the bandwidth of each pair in multi-pair VDSL2 access decreases sharply because of the increasing crosstalk. Crosstalk is the major issue degrading VDSL2 performance. Vectoring is developed to eliminate crosstalk on VDSL2 lines.
- Vectoring: detects crosstalk, compensates signals, and eliminates crosstalk on VDSL2 lines, providing a noiseless environment for VDSL2 lines to achieve the optimal VDSL2 performance. This technology not only maximizes copper line potential but also complies with the primary targets of broadcast access network construction.
- Multiple-input multiple-output (MIMO): combines vectoring and line bonding, increasing access rates by several times at the same distance using multiple twisted pairs. Two twisted pairs are widely used in site deployments.
- Super MIMO: developed based on MIMO, emulates (N - 1) pairs of virtual lines based on any N pairs of physical lines to achieve the transmission capabilities of (2N - 1) pairs of lines, further increasing access rates. Both MIMO and super MIMO apply to multi-pair drop access scenarios, such as commercial user access, mobile backhaul, and remote site backhaul.
- G.fast: applies to single-pair drop access scenarios and provides a 1 Gbit/s access rate (sum of downstream and upstream rates) within 100 m.
- 5GBB: latest copper line technology that provides an access rate of at least 5 Gbit/s within 50 m. Accordingly, the frequency range will be expanded to 500-800 MHz.

NOTE
Vectoring takes effect only on VDSL2 lines.

7.2 What Is Vectoring

Description

Vectoring is an ITU-T Recommendation G.993.5-compliant technology for improving VDSL2 line rates by eliminating far-end crosstalk (FEXT) on VDSL2 lines.

The vectoring technology uses vectored groups to jointly transmit signals in the downstream direction and receive signals in the upstream direction. This cancels FEXT on VDSL2 lines and increases multi-pair VDSL2 line rates. The crosstalk, a vector, on one VDSL2 line comes from the other lines in the same bundle. The central office (CO) device calculates the matrix based on the collected vector information and outputs vectored crosstalk cancellation signals to eliminate FEXT.

Vectoring System

Compared with the VDSL2 reference model, the vectoring reference model adds a vectoring control entity (VCE) as well as the interfaces between the VCE and a VDSL transceiver unit — office (VTU-O) and between the VCE and a management entity (ME), as red lines shown in Figure 7-2.

In a vectoring system, the access nodes (ANs) located at the CO, remote end, or other locations exchange data with multiple network terminals (NTs). Vectoring in all formats is implemented on the ANs in jointly transmitting signals (downstream vectoring) or receiving signals (upstream vectoring) over the lines in a vectored group. Therefore, all signals form a vector and each element in Figure 7-2 is a signal transmitted over lines.

In the vectoring system, the AN uses interface (ε-1-n) between one VTU-O interface (marked as VTU-O-1) and all other VTU-O interfaces (marked as VTU-O-n, where n is greater than or equal to 2 and less than or equal to the total number of lines in the vectored group) for jointly processing signals. Interface (ε-1-n) is used for jointly processing signals between lines 1 and N. Figure 7-2 shows only pair 1 in the vectored group.

NOTE

- An ME uses interface (ε-m) to manage the VCE. The VCE then uses interface (ε-c-n) to manage specified VTU-O interfaces in the vectored group. VTU-O interfaces correspond to vectoring lines in the vectored group.
- VTU-O interfaces use interface (ε-n1-n2) to exchange precoding data.

Figure 7-2 Vectoring reference model

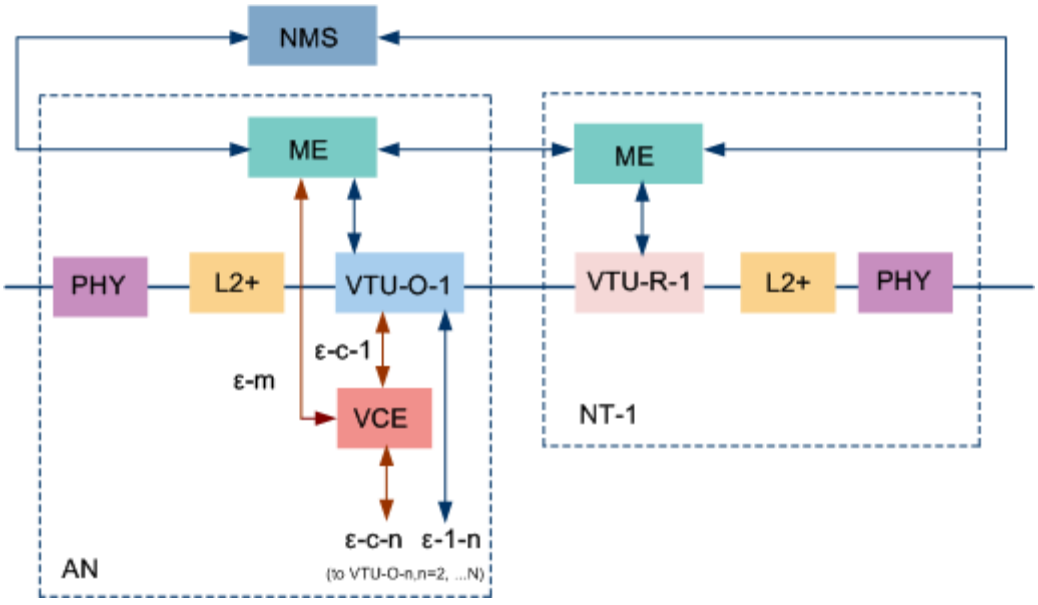


Table 7-1 Vectoring modules

Module	Description
ME	A management entity
PHY	Physical layer of the AN interface connected to the network and of the NT interface connected to customer premises (CPs)

Module	Description
L2+	Function implementation for the Layer-2 Ethernet and the upper-layer network where the ANs and NTs locate
VCE	Including a vector engine (VE) and a vector control unit (VCU) • The VE calculates matrices. • The VCU controls line activation, calculates and updates matrix coefficients, controls line joining to or leaving from a vectored group, and updates parameters in real time required for calculating crosstalk matrices, thereby ensuring the stability of the vectored group.

NOTE

Vectoring is different from the bonding of multiple pairs and can be used together with pair bonding. After vectoring is enabled on the lines with pairs bonded, these lines are bonded vectoring or MIMO DSL lines. Vectoring is developed based on unbonded pairs, but it also applies to bonded pairs.

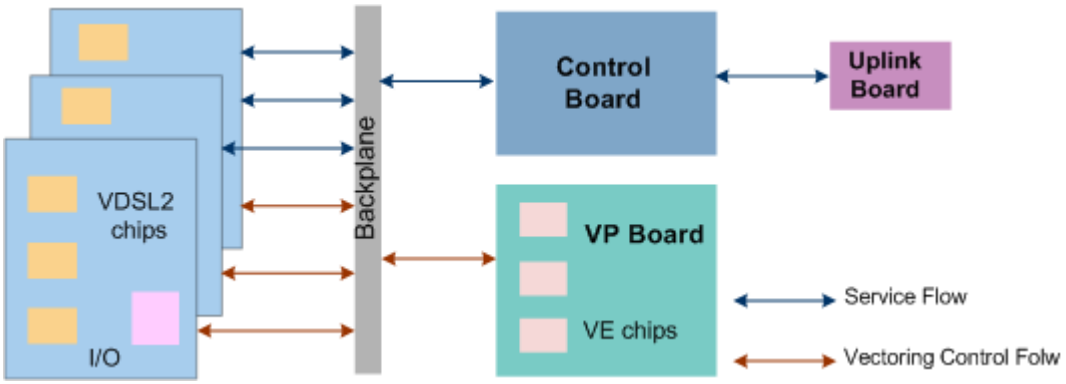
7.3 Vectoring Classifications

Vectoring is classified as system level vectoring (SLV) and node level vectoring (NLV) based on system architectures.

- SLV calculations are performed on a separate vectoring processing (VP) board. The VP board communicates with VDSL2 boards about crosstalk and crosstalk cancellation using a bus. SLV enables an access device to cancel crosstalk between the lines connected to multiple VDSL2 boards.
- NLV is developed based on SLV. Specifically, two SLV devices are connected using a high-speed cable. They work together to implement inter-device vectoring. NLV enables two connected SLV access devices to cancel crosstalk between the lines connected to the VDSL2 boards of the two devices.

SLV

Figure 7-3 SLV process flow

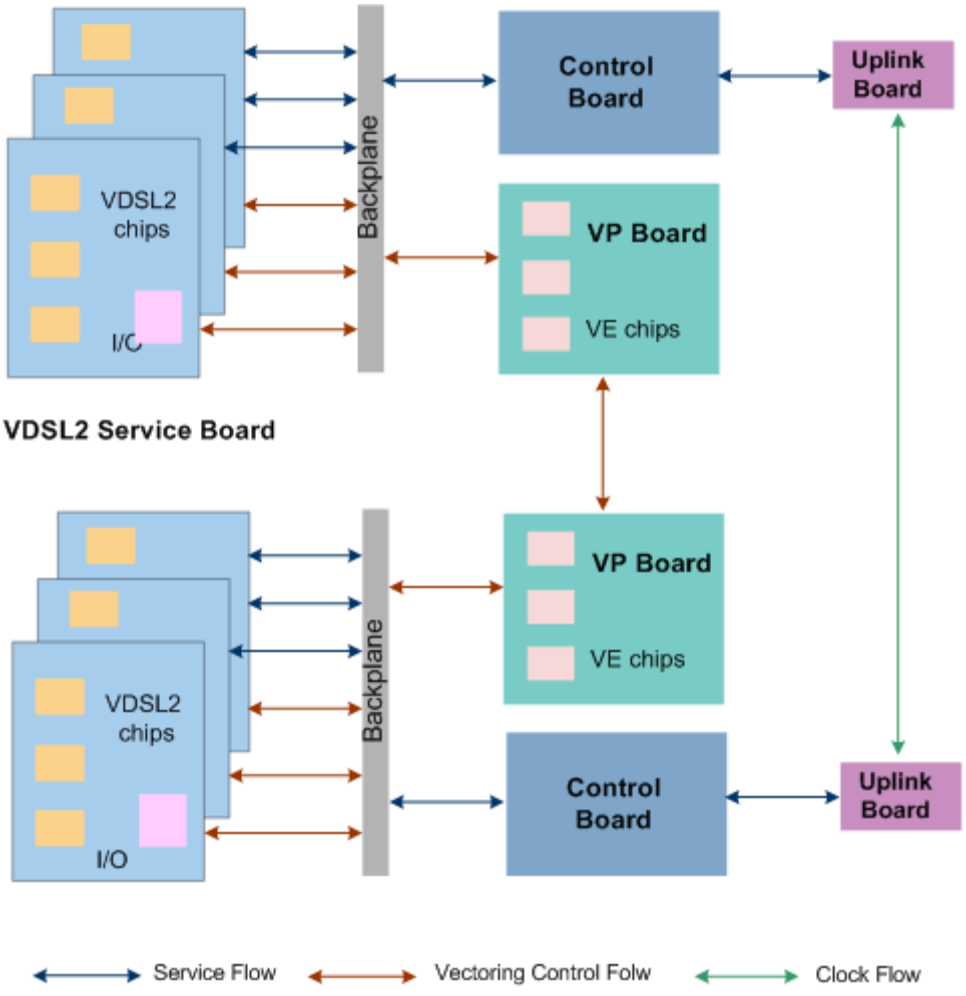


SLV characteristics are as follows:

- Vectoring processing chips are deployed on one VP board.
- An SLV system supports multiple vectoring service boards that totally connect to a maximum of 200 VDSL2 lines.
- The VP board communicates with vectoring service boards using a high-speed backplane bus.

NLV

Figure 7-4 NLV process flow



NLV characteristics are as follows:

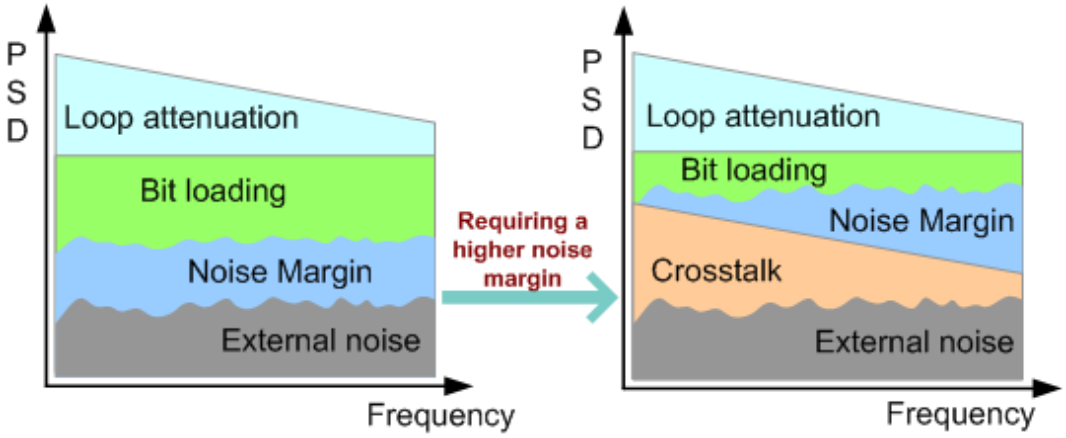
- Two SLV devices jointly implement vectoring.
- An NLV system supports multiple vectoring service boards that totally connect to up to 384 VDSL2 lines.
- The two SLV devices share data using a high-speed cable. In addition, clock synchronization must be enabled on both devices.

7.4 Vectoring Basic Concepts

7.4.1 Crosstalk

Crosstalk is the interference between pairs in one bundle when signals are coupled. VDSL2 uses high frequencies, which are more prone to crosstalk than low frequencies. Compared with the bandwidth of single-pair VDSL2 access, the bandwidth of each pair in multi-pair VDSL2 access decreases sharply because of the increasing crosstalk.

Figure 7-5 Impacts on lines caused by crosstalk

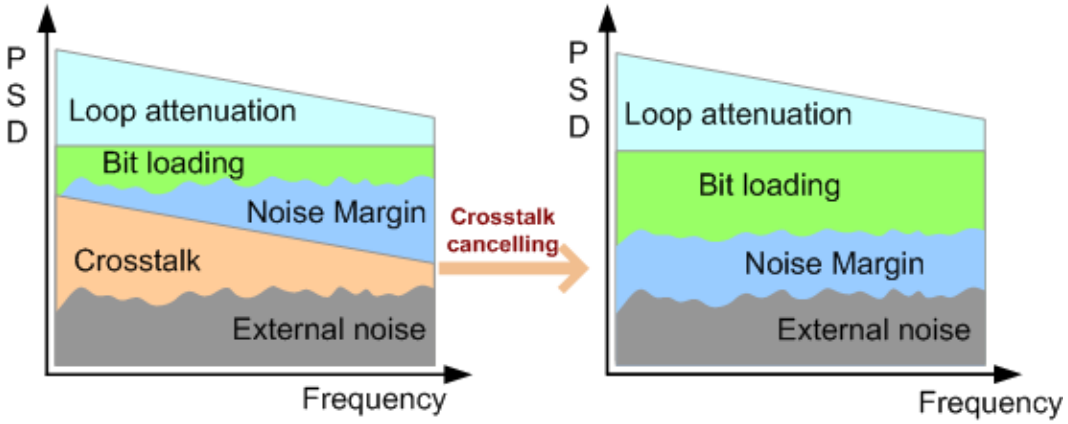


As shown in Figure 7-5, VDSL2 line rates are determined based on attenuation and noises on the lines. The louder noises on a line require a higher noise margin on the basis of the same transmission distance, thereby decreasing line rates (payloads).

VDSL2 line rates are also determined based on frequency bands. VDSL2 uses a high frequency band, 0-30 MHz, for short-distance transmission, where FEXT is the main noise interference. Based on tested data, crosstalk decreases VDSL2 line rates by 50%.

Line attenuation, external noises, and crosstalk together decrease VDSL2 line rates. If the line length, quality, and external environment cannot be improved, the only way to increase VDSL2 line rates is to cancel crosstalk.

Figure 7-6 Effects after crosstalk is canceled



Affected Scope of Inter-Line Crosstalk

A twisted pair cable can be of bundle, super unit, or basic unit type, containing at least 200 pairs, 100 pairs, or 25 pairs, respectively.

Crosstalk mainly occurs in basic units. The crosstalk between bundles and between super units slightly decreases VDSL2 line rates. The crosstalk in super units significantly decreases VDSL2 line rates.

Figure 7-7 Twisted pair cable catalog

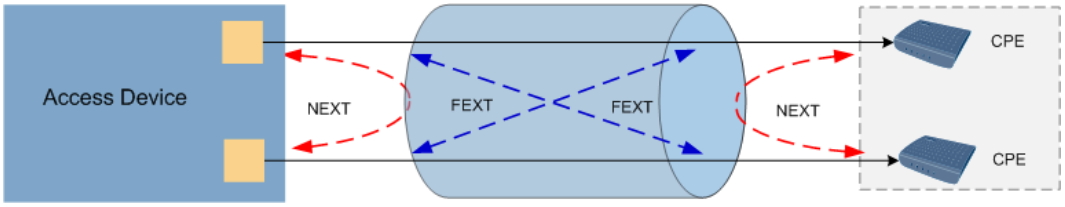
7.4.2 NEXT and FEXT

VDSL2 Crosstalk Classifications

VDSL2 crosstalk is classified as near-end crosstalk (NEXT) and far-end crosstalk (FEXT).

Figure 7-8 shows the crosstalk of the two types.

Figure 7-8 VDSL2 crosstalk



- In NEXT, TX signals are sent from the interfering pair, coupled to the interfered pair, and then sent to the near-end RX end of the interfered pair. For example, in a bundle of lines, when signals in the upstream direction of a line interfere with signals in the downstream direction of an adjacent line, or signals in the downstream direction of a line interfere with signals in the upstream direction of an adjacent line, NEXT occurs.
- In FEXT, TX signals are sent from the interfering pair, coupled to the interfered pair, and then sent along the interfered pair to the far-end RX end of the interfered pair. For example, in a bundle of lines, when signals in the upstream direction of a line interfere with signals in the upstream direction of an adjacent line, or signals in the downstream direction of a line interfere with signals in the downstream direction of an adjacent line, FEXT occurs.

In other words, NEXT is interference between upstream signals and downstream signals of different pairs, and FEXT is interference between upstream signals or between downstream signals of different pairs.

How Can We Eliminate NEXT and FEXT

- VDSL2 uses the frequency division multiplexing (FDM) technology to transmit data. Therefore, TX signals of the interfering pair and RX signals of the interfered pair use different frequencies. Therefore, the impact of NEXT can be eliminated or mitigated using a filter.
- TX signals of the interfering pair cannot be eliminated using a filter because these signals use the same frequency band as the RX signals of the interfered pair. In addition, VDSL2 uses a high frequency band (up to 30 MHz) for short-distance transmission (usually within 1.2 km). As a result, FEXT has a more severe effect on VDSL2 than on other DSL access modes. **Therefore, FEXT is the main factor of degrading VDSL2 performance. To eliminate FEXT, the ITU-T Recommendation promoted G.993.5-compliant vectoring.**

NOTE

- To eliminate or mitigate crosstalk, the DSL industry promoted a series of techniques totally called the dynamic spectrum management (DSM) technology. The DSM technology involves four stages, level 0 through level 3 stages. At level 0 through level 2 stages, the AN manages the spectra of the TX signals of single- or multi-DSL pairs, which eliminates FEXT only to a certain extend. To completely cancel FEXT on VDSL2 lines, the ITU-T Recommendation launched vectoring. Vectoring uses vectors to cancel FEXT on VDSL2 lines, thereby significantly improving the bandwidth and performance of multi-pair VDSL2 lines. Therefore, vectoring is also called level-3 DSM.
- Vectoring can significantly eliminate only FEXT.

7.4.3 Vectoring CPE Classifications

Customer premises equipment (CPE) devices are classified as vectoring CPEs, vectoring friendly CPEs, and legacy CPEs based on CPE statuses in supporting vectoring.

Table 7-2 Vectoring CPE classifications

CPE Type	Functions Supported in the Upstream Direction	Functions Supported in the Downstream Direction	Remarks
Vectoring CPE	Is able to send test signals to the CO device.	Is able to receive test signals from the CO device and return crosstalk information to the CO device.	• Vectoring CPEs can be activated in G.993.5 mode. • Vectoring can improve the performance of vectoring CPEs.
Vectoring friendly CPE	Is able to send test signals to the CO device.	Is able to receive test signals from the CO device.	• Vectoring friendly CPEs do not support vectoring, but they accept the vectoring process flow. • Vectoring cannot improve the performance of vectoring friendly CPEs but can mitigate the crosstalk of such CPEs on other lines in the same vectored group.
Legacy CPE	Is not able to send test signals to the CO device.	Is able to receive test signals from the CO device.	• Legacy CPEs can only be activated in G.993.2 (VDSL2) mode. • Vectoring can neither improve the performance of legacy CPEs nor mitigate the crosstalk of such CPEs on other lines in the same vectored group. Legacy CPEs are processed using the common VDSL2 process flow. The activated legacy CPEs degrade the vectoring performance in the same vectored group.

Figure 7-9 Vectoring CPE networking

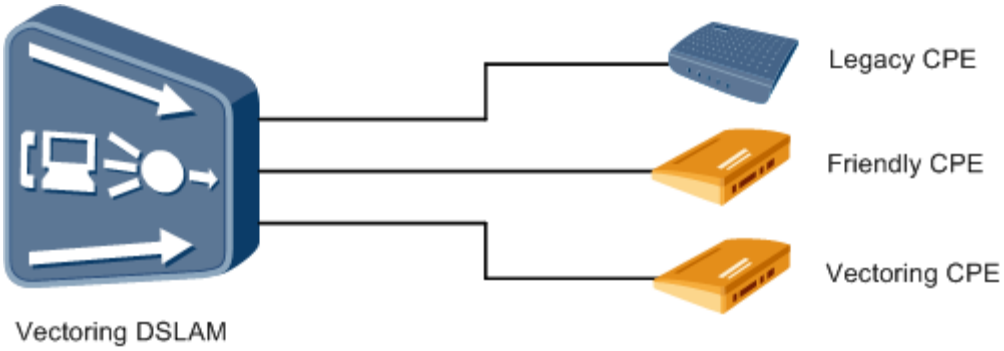
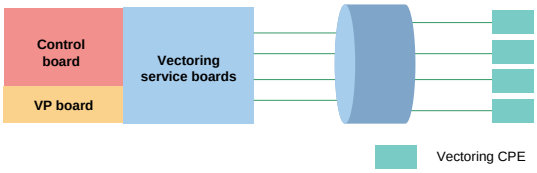
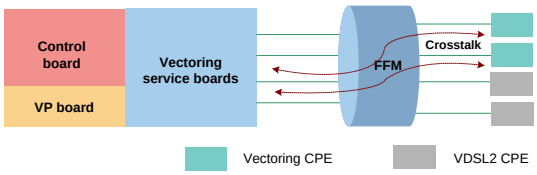


Table 7-3 CPE networking characteristics

CPE Activation Flow	Networking Characteristics	Description
G.993.5 or G.993.5 friendly (vectoring and vectoring friendly CPEs)		• Ideally, both service boards and CPEs support vectoring, which maximally increase VDSL2 line rates. • No action is required and these CPEs can be activated normally.
G.993.2 (legacy CPEs)		• Both vectoring and legacy CPEs are deployed in the same network. • Legacy CPEs cannot send crosstalk information to the CO device and line crosstalk cannot be eliminated. Therefore, VDSL2 line rates cannot be improved.

NOTE

Legacy CPEs significantly degrade vectoring performance and the impact can be decreased by configuring legacy CPE activation policies. For details, see Activation Policies for Legacy CPEs.

7.5 Vectoring Applications

7.5.1 Site Planning

Based on carriers' live network deployment, vectoring sites are classified as new vectoring sites and reconstructed vectoring sites.

Table 7-4 Site planning scenarios

Scenario	Description
New vectoring site	A new site meets one the following requirements: • Is newly constructed. • Services and users are brand new, although the site is reconstructed. • The site is upgraded to support new vectoring services and users. • Although original services and users are in service on the site, these services and users do not or slightly affect new vectoring services and users.
Reconstructed vectoring site	In a reconstructed site, the original services and users interfere with new vectoring services and users and the interference cannot be eliminated. Consider the following factors before enabling vectoring: • Existing subscriber lines do not severely degrade vectoring performance. • The reconstructed site supports the coexistence of ADSL2+, VDSL2, and vectoring services. • ADSL, ADSL2, and ADSL2+ services do not severely degrade vectoring performance. In addition, after vectoring is enabled, the performance of original DSL lines is not degraded and the performance of entire system is improved.

Figure 7-10 Site planning diagram

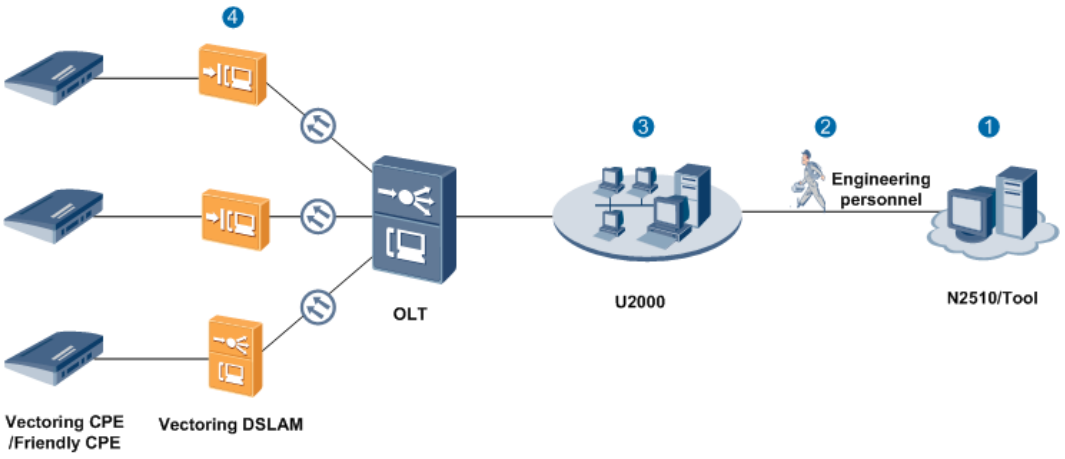


Table 7-5 Site planning process

No.	Operations for New Sites	Operations for Reconstructed Sites
1	**N2510 pre-evaluation:** • Use the pre-evaluation function provided by the N2510 to evaluate the maximum and minimum rates for provisioning services. • Evaluate vectoring rates supported by lines with various lengths, especially rate increase after vectoring is enabled. In addition, use desired tools to provide vectoring evaluation reports based on CO cable types and service planning.	Perform the same operations as those required for new sites.
2	**Engineering personnel:** Ensure that the hardware and software of the devices to be deployed support vectoring. Then, deploy the devices and connect cables.	Perform the same operations as those required for new sites.
3	**U2000:** • If VDSL2 service boards are securely connected, configure vectoring and provision services using the U2000. • If VDSL2 service boards are not securely connected, pre-deploy vectoring using the U2000. When vectoring needs to be provisioned, securely connect the VDSL2 service boards. Then, the pre-deployed vectoring automatically takes effect on these VDSL2 service boards.	**U2000:** • Upgrade the U2000 to the desired version. Ensure that the U2000 supports existing xDSL profiles, preventing significant adjustment on the operations support system (OSS). TR165 profiles are recommended. • If VDSL2 service boards are securely connected, configure vectoring and provision services using the U2000. • If VDSL2 service boards are not securely connected, pre-deploy vectoring using the U2000. When vectoring needs to be provisioned, securely connect the VDSL2 service boards. Then, the pre-deployed vectoring automatically takes effect on these VDSL2 service boards.
4	**Vectoring DSLAM:** • Ensure that the board supporting vectoring in hardware is available.	**Vectoring DSLAM:** • Ensure that the board supporting vectoring in hardware is available. (Replace hardware if required based on the mapping between the

No.	Operations for New Sites	Operations for Reconstructed Sites
	• If the U2000 is available, issue vectoring configurations using a northbound interface (NBI). • If the U2000 is not available, configure vectoring by running a command or script.	control board, VP board, backplane, and VDSL2 service boards. If the DSLAM is an integrated device, replace the entire device to a vectoring-supported DSLAM.) • If the U2000 is available, issue vectoring configurations using an NBI. • If the U2000 is not available, configure vectoring by running a command or script.
• **Planning for new sites**: The operations required for provisioning the vectoring service are the same as those required for provisioning a common service. Ensure that cables are connected and data is planned before provisioning the vectoring service. In the new site, use the vectoring-supported components. No impact on new lines brought by original lines is involved. • **Planning for reconstructed sites**: Operations required for planning reconstructed sites are more complex than those required for planning new sites. Consider the deployment and running status of live network devices and VDSL2 lines at planning phase. In addition, consider the impact on vectoring performance brought by the live network devices and VDSL2 lines. • **Requirements on CPEs**: The CPEs must be vectoring or vectoring friendly CPEs. If the vectored group contains legacy CPEs, do not activate these legacy CPEs after enabling vectoring, preventing the legacy CPEs from generating crosstalk on other CPEs in this vectored group. Therefore, upgrade legacy CPEs to vectoring-supported CPEs before enabling vectoring.		

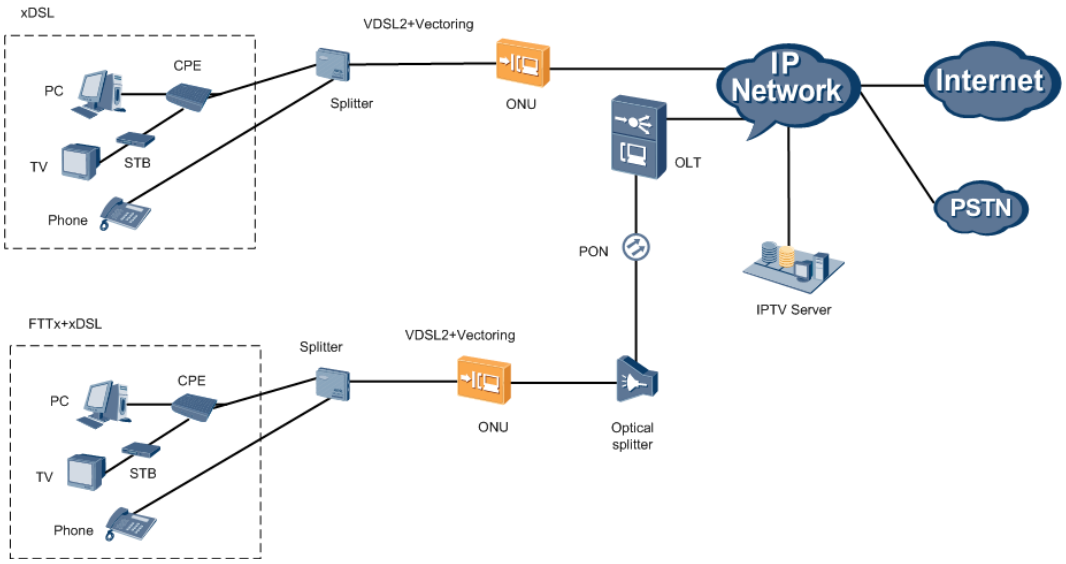
NOTE

Cross-connected cable sites are a special scenario, where the crosstalk cannot be eliminated only when a site shares a cable with other sites, regardless whether the sites are new or reconstructed ones. In carrier markets, some carriers rent an entire or part line to other carriers and therefore cross-connected cable sites are common on the live network. For a new or reconstructed site, if this site shares cables with another CO or remote terminal (RT) site, consider the impact on vectoring performance brought by original subscriber lines before provisioning vectoring configurations to this site. If vectoring cables are used with traditional VDSL2 cables, consider the impact on vectoring-supported lines brought by the lines not supporting vectoring.

7.5.2 Network Application

Vectoring, a new generation of technology for promoting line performance, is compatible with existing DSL technologies, including retransmission (G.INP), bonding, network time reference (NTR), seamless rate adaptation (SRA), and bit swap (BS). With these technologies, vectoring can be flexibly used in various scenarios, such as residential user access, commercial user access, mobile base station backhaul, and remote access site backhaul.

Figure 7-11 Typical vectoring networking



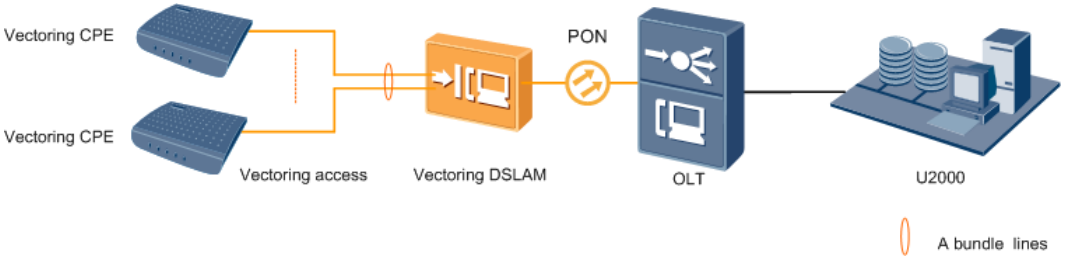
The following two scenarios provide information about how vectoring is used in the two networks:

NOTE

Vectoring has the same requirements on lines, connectors, and line sequence as common VDSL2 technologies.

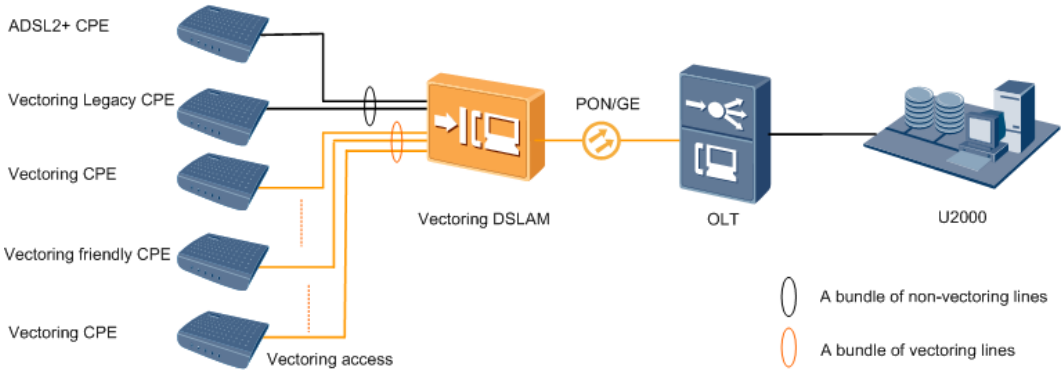
Scenario 1: used for FTTB networks. In this scenario, there are 1-2 bundles of VDSL2 lines, covering less than 50 access users within a 300-meter reach.

Figure 7-12 FTTB networks scenario



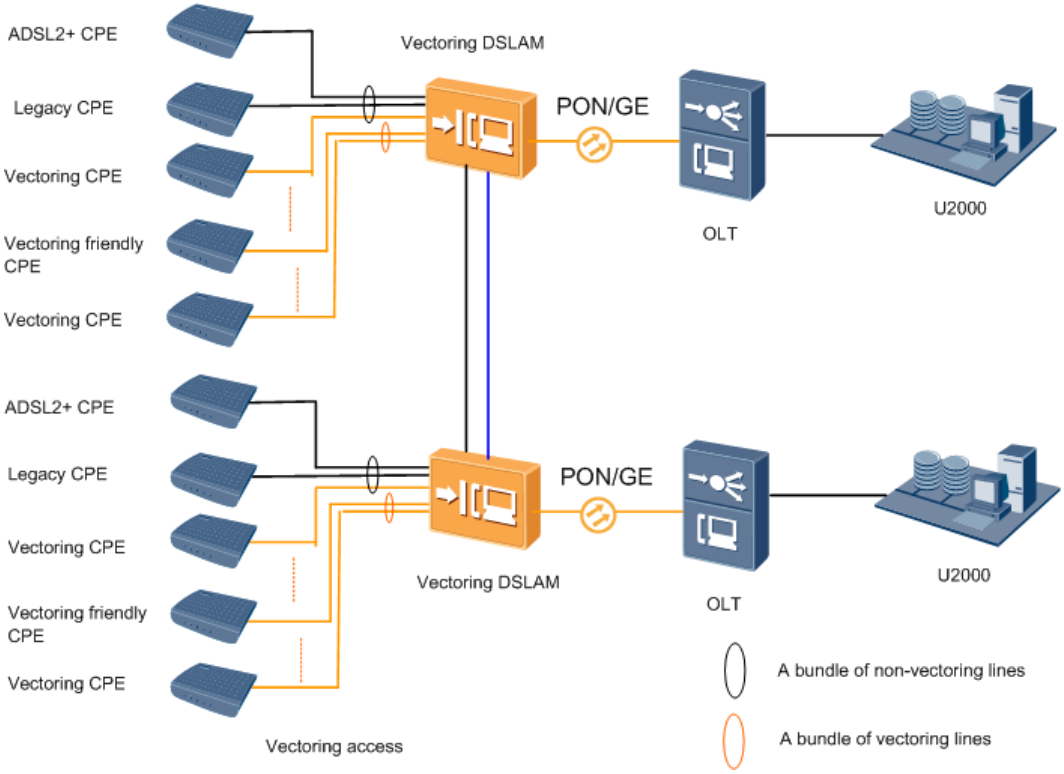
Scenario 2: mainstream scenario for FTTC networks. In this scenario, there are 1-8 bundles of VDSL2 lines, covering 100-300 access users over a 300-800-meter reach.

Figure 7-13 FTTC networks scenario



In NLV scenario, the two SLV devices share data using a cable. Two SLV devices jointly implement vectoring.

Figure 7-14 NLV scenario



NOTE

- Connect the GE ports and CXP ports on power boards of the two SLV devices, respectively, to form an NLV system, which cancels crosstalk on vectoring lines connected to VDSL2 boards of the two devices.
- Use a network cable to connect the GE ports on two SLV devices.
- Use a CXP high-speed cable to connect the CXP ports on two SLV devices. One end of the CXP high-speed cable connects to the VEI port on the cascading SLV device. The other end of the cable connects to the VEI port on the cascaded SLV device.

7.5.3 Vectoring Engineering Precautions

Configuration Requirements on Vectoring Lines

- In vectoring algorithm, the system requires that the upstream frequencies and downstream frequencies are separate from each other. Therefore, frequencies must be planned globally and then the Vectoring is enabled globally. After Vectoring is enabled globally, the system activates the port according to its profile's frequency planning and the compatibility between the frequency planning and global vectoring frequency planning.
  - If the profile's frequency planning is incompatible with global vectoring frequency planning, the port cannot be activated in G.993.5/G.993.2. If the line profile of the port contains other transfer modes expect for G.993.2, the port can be activated in non-VDSL mode.
  - If the profile's frequency planning is compatible with global Vectoring frequency planning and a Vectoring legacy CPE is used, whether the CPE is allowed to be activated is determined by the policy for activating the legacy CPE.
- The maximum rate in the line profile bound to a port is the target value for vectoring rate improvement. Carriers need to plan the line profile based on services.
- The maximum rate in the line profile bound to a port is the target value for vectoring rate improvement. Carriers need to plan the line profile based on services.
- All VDSL2 lines in one bundle must belong to one vectored group and support vectoring. Do not enable vectoring if only some VDSL2 lines support vectoring. The reason is that the VDSL2 lines not supporting vectoring will cause external noises that cannot be canceled for the VDSL2 lines supporting vectoring, degrading vectoring performance or even causing vectoring to fail to take effect.

Vectoring Application Notes

- Vectoring port enabling is controlled by licenses.
- VDSL2 lines must be shorter than 1 km.
- VDSL2 lines using profile 30a do not support vectoring.
- **Impact on vectoring caused by VDSL2 lines:** A non-vectoring VDSL2 line affects vectoring performance significantly at a site. Therefore, do not use vectoring and non-vectoring VDSL2 lines at the same site.
- **Impact on vectoring caused by ADSL2+ lines** An ADSL2+ line affects vectoring performance slightly at a site or CO. Therefore, vectoring and ADSL2+ lines can be used at the same site.
- All VDSL2 CEPs are recommended to support vectoring.

7.5.4 Vectoring Hardware

Table 7-6 Vectoring Hardware

Product	Board Type	Board Name	Remarks	Corresponding Outdoor Cabinet	Terminals
MA5603T	**Control board**	SCUB SCUN	Supports SLV only.	S300	Vectoring can be implemented

Product	Board Type	Board Name	Remarks	Corresponding Outdoor Cabinet	Terminals
		SCUK			on VDSL2 lines only when their connected terminals support vectoring. Vectoring-supported Huawei terminals include the HG612, HG622, HG630, and HG658. For details about the version of these terminals in supporting vectoring, see the product documents of these terminals.
	Backplane	H802MABO	None		
	VP board	H806VPEA	Installed in slot 12 fixedly.		
	VDSL2 board	• H80BVCMM • H80DCCPE • H80DVCPD • H80DVCPE • H80DVCPM	• The H80BVCMM board is a 48-channel VDSL2 over POTS access service board. • The H80DCCPE board is a 64-channel VDSL2&POTS Combo Board with built-in splitter. • The H80DVCPD board is a 64-channel VDSL2 over POTS access service board. • The H80DVCPE board is a 64-channel VDSL2 over POTS access service board, equipped with a built-in splitter. • The H80DVCPM board is a 64-channel VDSL2 over POTS access		

Product	Board Type	Board Name	Remarks	Corresponding Outdoor Cabinet	Terminals
			service board.		
MA5600T	**Control board**	SCUB SCUN SCUK	Supports SLV only.	N/A	
	Backplane	H802MABC H803MABC	None		
	VP board	H806VPGA	Consistently installed in slot 8 and slot 11.		
	VDSL2 board	• H80BVCMM • H80DCCPE • H80DVCPD • H80DVCPE • H80DVCPM	• The H80BVCMM board is a 48-channel VDSL2 over POTS access service board. • The H80DCCPE board is a 64-channel VDSL2&POTS Combo Board with built-in splitter. • The H80DVCPD board is a 64-channel VDSL2 over POTS access service board. • The H80DVCPE board is a 64-channel VDSL2 over POTS access service board, equipped		

Product	Board Type	Board Name	Remarks	Corresponding Outdoor Cabinet	Terminals
			with a built-in splitter. • The H80DVCPM board is a 64-channel VDSL2 over POTS access service board.		
MA5623AR	N/A	N/A	The MA5623AR extended subrack can be considered as the extension service board for the main subrack. The main subrack manages the MA5623AR extended subrack in the same way as it manages its service boards. The MA5623AR extended subrack provides the same functions as the VDSL2 board of the main subrack.	N/A	
MA5616	**Control board**	H831CCUE	Supports SLV and NLV.	S200/S100 **NOTE** The S200 cabinet is recommended because it supports a maximum of 192 lines. The S100 cabinet supports only 96 lines due to the limitation of	
	Daughter board	UP2CA/UP2AA	None		
	Backplane	H831MABB	None		
	VP board	H836VPBA H836V	• H836VPBA: A daughter board for		

Product	Board Type	Board Name	Remarks	Corresponding Outdoor Cabinet	Terminals
		PDA	SLV; attached to the power board. • H836VPDA: A daughter board for NLV; attached to the power board.	heat dissipation.	
	VDSL2 board	H83BVCMM H83BVCLE H83BVCLF	• H83BVCMM: A 48-channel VDSL2 access service board. • H83BVCLE/H83BVCLF: 32-channel VDSL2 access service board.		
	Power board	H831PAVDA H832PDVAA H832PDNAA	• H831PAVDA: An AC power board for SLV. • H832PDVAA: A DC power board for SLV. • H832PDNAA: A DC power board for NLV.		
MA5622A	**Control board**	HS22CCVB	Supports SLV only.	N/A	
MA5623A	**Control board**	HS22CCVW	Supports SLV only.	N/A	
MA5611S	**Integrated device**	N/A	Supports SLV only.	N/A	
MA5811S	**Integrated device**	N/A	Supports SLV only.	N/A	

NOTE

Before enabling vectoring, check whether your device supports vectoring in hardware.

- To query the information about a control board, service board, VP board, daughter board, or power board, run the **display board** command.
- To query the information about a backplane, run the **display version backplane** command.

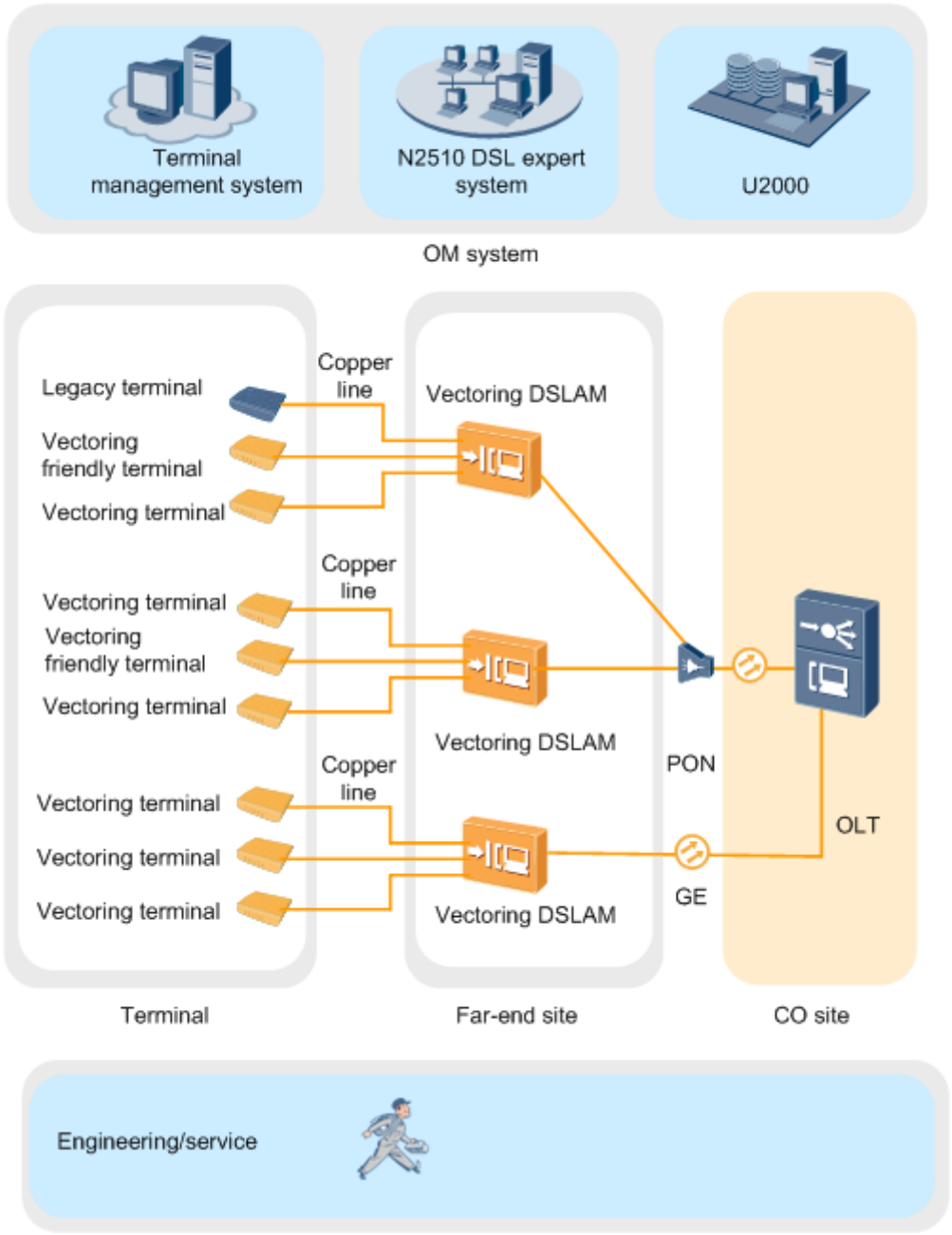
7.6 Vectoring Implementation Principles

This section describes how is vectoring implemented from the aspect of system components, principles, flow, and key techniques.

7.6.1 System Architecture

Huawei Vectoring solution incorporates CO devices, terminals, NMS, terminal management system, DSL expert system, and corresponding engineering and service solution. It can be deployed in batches and is ease of management and control. Figure 7-15 shows architecture of the Vectoring solution.

Figure 7-15 Architecture of the Vectoring solution



Components in Figure 7-15 provide their respective functions:

- Vectoring DSLAM: Includes series Vectoring DSLAMs with different capacities, to suit various site scales and deployment scenarios. A Vectoring DSLAM needs to support traditional DSL technologies (such as VDSL2+, ADSL2+, and ADSL), plug-and-play of different types of terminals, and smooth evolution to Vectoring. Based on system architecture, Vectoring DSLAM can be classified to two types: System level vectoring (SLV), and Node level vectoring (NLV).
- Vectoring terminal: Includes terminals that fully support Vectoring and Vectoring friendly terminals. Generally, VDSL2 terminals deployed on live network can become Vectoring terminals or Vectoring friendly terminals after software upgrades. Legacy CPE

can be activated only by using a common VDSL2 standard (ITU-T Recommendation G.993.2). The crosstalk impact of such CPEs cannot be alleviated in a vectoring system, and such CPEs will impair performance of the vectoring system.

NOTE

The system generates the event if the following conditions are met:

- The vectoring function is enabled.
- The transmission mode of the port is 993.5.
- The peer customer premises equipment (CPE) connected to the port does not support the vectoring function or the vectoring friendly function.
- The port cannot be activated in G.993.5 mode when the legacy CPE policy does not allow port activation.

- U2000: Provides graphic interface. It supports Vectoring service provisioning and configration, and plans and controls the schedule of Vectoring service provisioning.
- N2510 DSL expert system: Monitors DSL quality, evaluates and optimizes DSL performance, and diagnoses copper line faults at a network or site level. To support Vectoring deployment and OM, the DSL expert system needs to provide Vectoring-specific functions, such as pre-evaluating Vectoring performance, coordinating coexistence of Vectoring and other DSL lines, processing combined application of Vectoring and other DSL features, and preventing and processing Vectoring abnormalities. Providing these functions, the DSL expert system helps achieve Vectoring capabilities that the Vectoring equipment or NMS cannot provide independently.
- Terminal management system: Manages, upgrades, and maintains terminals in a centralized manner. In an ideal environment for Vectoring deployment, all terminals on the entire network (or at least on the entire site) support Vectoring. Therefore, it is necessary to use the terminal management system to upgrade VDSL2 terminals on live networks beforehand.
- Engineering/service: Provides services such as network panning, equipment migration, equipment upgrade, data planning, and data migration based on the Vectoring evolution/deployment scenarios and equipment models/versions on live networks.

7.6.2 Vectoring Principles

Vectoring uses pre-coder and canceller to cancel inter-VDSL2 line crosstalk in downstream and upstream directions, respectively.

Pre-coder Applied in the Downstream Direction

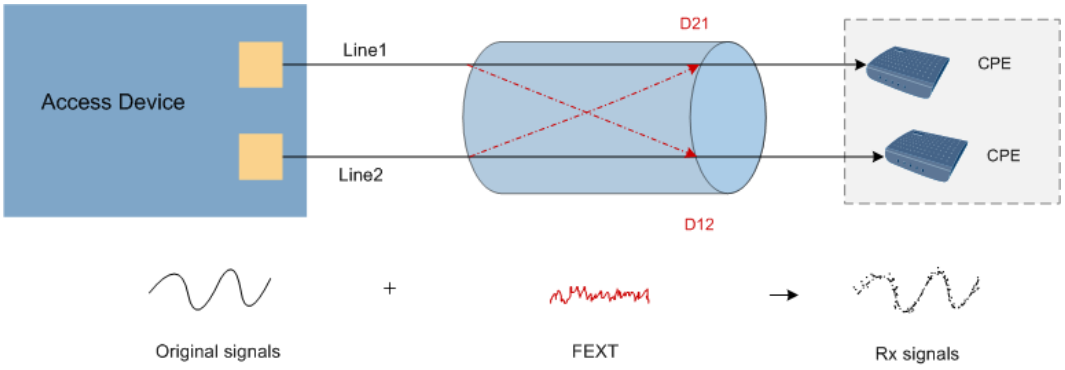
NOTE

This section uses two pairs as an example to describe how is pre-coder implemented.

When vectoring is disabled

As shown in Figure 7-16, VDSL2 lines 1 and 2 belong to one bundle. The two lines cause crosstalk in the downstream direction, degrading VDSL2 line performance. In Figure 7-16, crosstalk signals from line 1 are identified as D12 and crosstalk signals from line 2 are identified as D21.

Figure 7-16 Crosstalk on VDSL2 lines in the downstream direction when vectoring is disabled

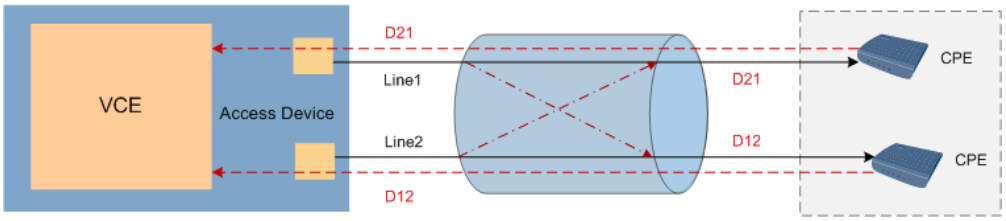


When vectoring is enabled

Pre-coder is implemented as follows:

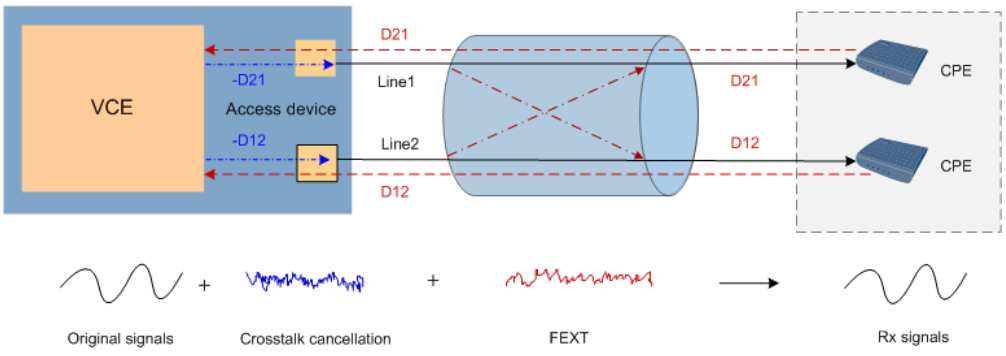
1. The VCE issues test data to all lines in real time. When receiving the test data, the CPEs send collected ESs to the VCE through data channels, as shown in Figure 7-17.

Figure 7-17 ES data collection on VDSL2 lines in the downstream direction



2. After receiving the ESs from the CPEs, the VCE calculates the vectoring matrix and derives cancellation signals -D21 and -D12 (reverse crosstalk signals) for D21 and D12, respectively.
3. The VCE superimposes the -D21 cancellation signals to line 1 and -D12 cancellation signals to line 2 in the downstream direction to cancel crosstalk. After the operation, the CPEs receive original signals that are not affected by crosstalk. In this way, VDSL2 line performance is improved significantly, as shown in Figure 7-18.

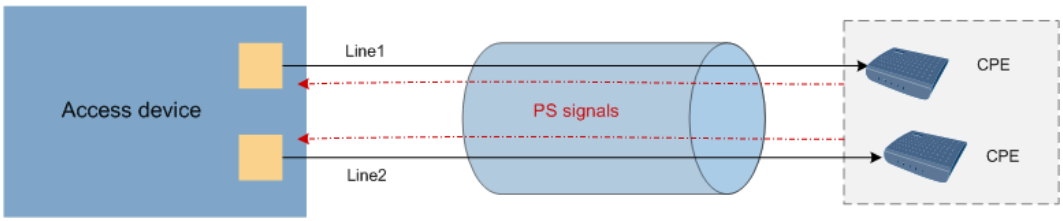
Figure 7-18 Crosstalk cancellation on VDSL2 lines in the downstream direction



Canceller Applied in the Upstream Direction

The principles of canceller are similar to those of pre-coder and therefore will not be described in this document. The only difference between the two techniques is as follows: In canceller, CPEs send test data (PS signals) to the VCE and the VCE calculates crosstalk information based on the received test data; in pre-coder, the VCE sends test data to CPEs, the CPEs send ESs back to the VCE, and the VCE calculates crosstalk information based on the ESs.

Figure 7-19 PS signal transmission on VDSL2 lines in the upstream direction



NOTE

- In the preceding description, the test data includes test signals and PS signals. The signals sent by the VCE for the first time are test signals, which are notification signals. The signals sent by the VCE for the second time are PS signals. The CPEs compare received PS signals with standard PS signals and send compared results (ES signals) back to the VCE.
- Crosstalk cancellation effects vary depending on CPE capabilities. If a CPE supports the sending of ESs to the VCE, the VCE can calculate the crosstalk cancellation coefficient required by the CPE port connected to the VCE. In this way, the crosstalk on this port caused by other ports can be canceled, thereby improving the rate of this port. If a CPE does not support the sending of ESs to the VCE, the CPE port connected to the VCE reports an ES loss alarm to the VCE. Then, the crosstalk on this port caused by other ports cannot be canceled and the rate of this port cannot be improved.
- ESs stand for the crosstalk on a port caused by other ports. If a port does not support the sending of ESs to the VCE, the rate of only this port cannot be increased.

7.6.3 Vectoring Flows

Vectoring flows include join in, tracking, and disorderly leaving event (DLE)/orderly leaving event (OLE).

Figure 7-20 shows the conversion between vectoring flows.

Figure 7-20 Conversion between vectoring flows

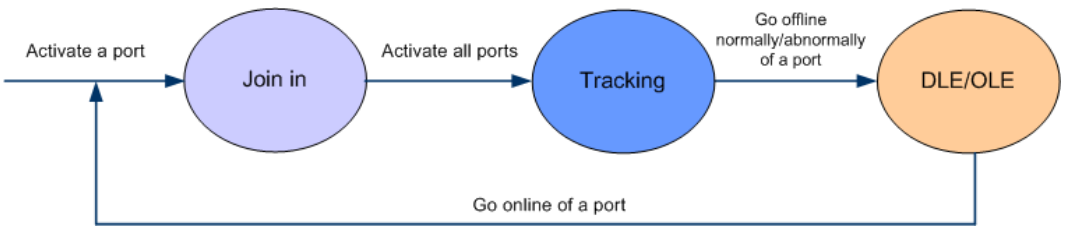


Table 7-7 Flow description

Flow name	Description

Flow name	Description
Join in	In join in process, one or multiple ports are concurrently activated. The process of activating a port in the vectoring system is different from that of activating a common VDSL2 port. Specifically, when a new port needs to be activated, the AN calculates the crosstalk cancellation coefficient based on the configurations and running status of the port to be activated and all other ports in the same vectored group. Then, the AN issues calculated line parameters to ports.
Tracking	In tracking process, the AN updates the pre-coder or canceller coefficient when all lines are activated. The AN calculates the pre-coder or canceller coefficient at initialization phase, which is incomplete for all channels. Therefore, the AN periodically calculates the pre-coder or canceller coefficient, not only providing a proper coefficient for all channels but also adapting to channel changes.
DLE/OLE	In DLE process, a port goes offline without any negotiation between the TX and RX ends. If DLE occurs, crosstalk channels may change, which reduces the SNR of other lines or even causes ports connected to other lines to go offline. DLE significantly degrades vectoring performance. To minimize DLE impact on vectoring performance, the AN must immediately take measures after detecting a port where DLE occurs. The measures include terminating the join in or tracking task and disabling signal transmission on DLE lines. In OLE process, a port goes offline after a negotiation between the TX and RX ends.

NOTE

- A DLE may be caused by many reasons. For example, the VTU-R is powered off; the cable is disconnected from the VTU-O or VTU-R; the line is cut.
- The impact on other lines caused by a DLE line is that the changed crosstalk channel does not match the channel after vectoring training and other lines require a period of time to adapt to the crosstalk channel change. Although this adaptation time is not long, the adaptation degrades vectoring performance. For example, bit errors occur. If DLEs concurrently occur on multiple DLE lines, channel environment deteriorates so sharply that vectoring ports may go offline.

Vectoring also involves grouping, which can be performed manually or automatically.

- Manual grouping: Users manually group cables based on cable status.
- Automatic grouping: The AN uses an intelligent algorithm to automatically group all lines based on the pre-coder or canceller coefficient.

NOTE

- If the grouping is specific in a vectoring system, group cables based on cable connections.
- The default grouping mode used by Huawei is that the AN automatically adds all lines to default vectoring group 1.

7.6.4 Key Vectoring Techniques

A series of key vectoring techniques are applied to improve bandwidths and stabilities of VDSL2 lines.

Vectoring Status

Enabled vectoring allows an AN to jointly process downstream and upstream signals to eliminate FEXT, thereby significantly improving VDSL2 line performance.

NOTE

- To enable vectoring, run the **xdsl vectoring** command.
- Enabled vectoring takes effect only when the VP board is functional.
- **Enabling and disabling vectoring interrupt services on ports in a vectored group.**

Vectoring applies to FTTB and FTTC scenarios, where subscriber line lengths must be shorter than 1000 m. The following section uses a vectoring test on 24 subscriber lines as an example to describe the relationships between downstream and upstream attainable rates and transmission distances.

NOTE

Downstream and upstream attainable rates provided in the following figures are obtained in ideal network conditions and for reference only. They vary depending on network planning and hardware.

Figure 7-21 Relationship between downstream attainable rates and transmission distances

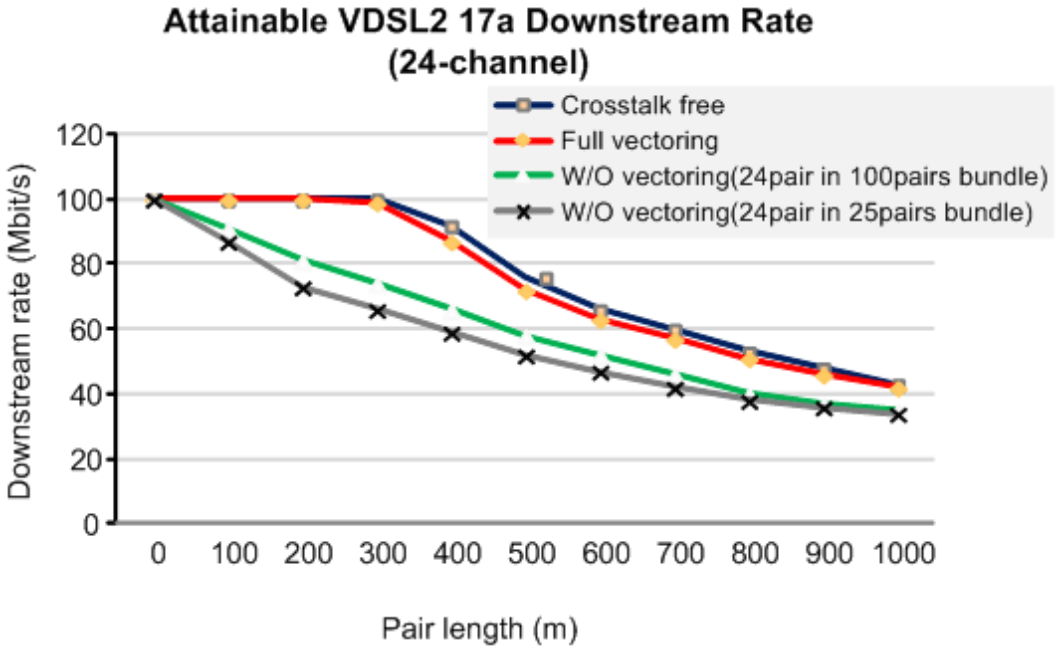
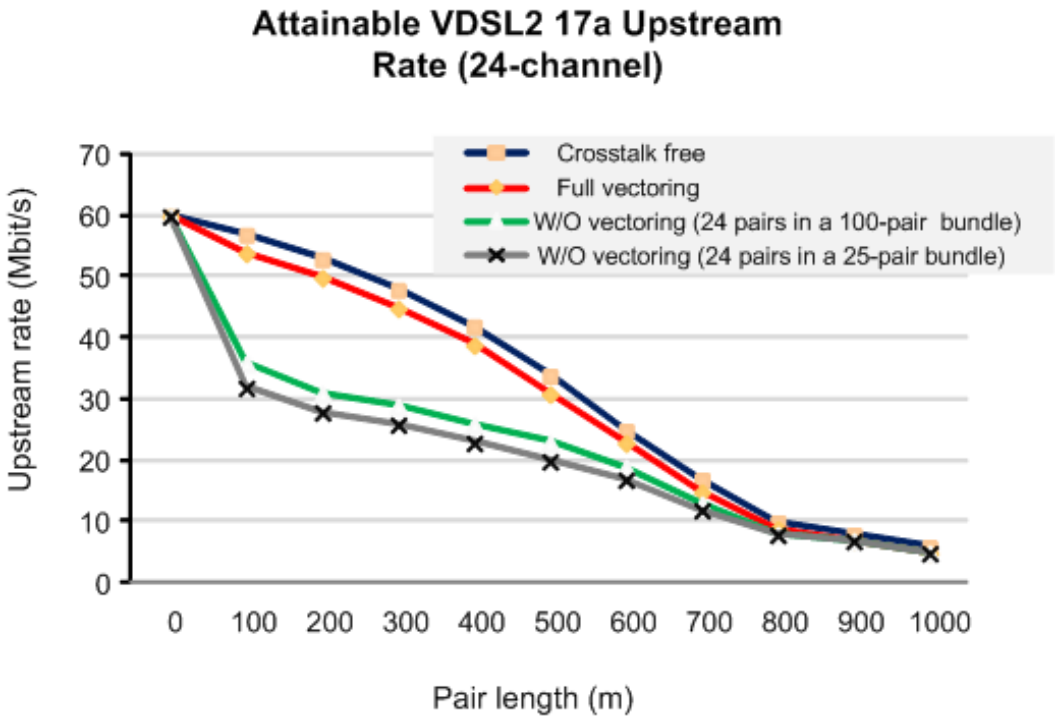


Figure 7-22 Relationship between upstream attainable rates and transmission distances



Based on the preceding figures, vectoring significantly takes effect if the transmission distance ranges from 200 m to 800 m. It slightly takes effect if the transmission distance is longer than 1000 m.

- Compared with common VDSL2 techniques, vectoring increases the rate of a single VDSL2 line by 50%-90% within 500 m, reaching 95% of the theoretical crosstalk-free VDSL2 rate.
- Increased VDSL2 line rates support more service types for users, such as HDTV.

Band Plans

Before switching services on a port from VDSL2 only to vectoring for increasing rates, check whether the vectoring band plan is compatible with the band plan specified in the VDSL2 line profile bound to this port. Ensure that the no frequency bands are overlapped in both downstream and upstream directions. Otherwise, this port cannot be activated.

Plan data based on the relationships between line profiles and global band plans before switching services, as shown in Table 7-8. For details about limit power spectrum density (PSD) masks, see 6.3.7 Limit PSD Mask.

NOTE
G.993.2 standards are continuously updating. Therefore, the data listed in Table 7-8 may be not the latest and is for reference only.

Table 7-8 Relationships between line profiles and global band plans

Short Name	Limit PSD Mask (Long Name)	US0 Type A/B/M	Compatible Band Plan Type

Short Name	Limit PSD Mask (Long Name)	US0 Type A/B/M	Compatible Band Plan Type
B7-1	997-M1c-A-7	A	997E
B7-2	997-M1x-M-8	M	997E
B7-3	997-M1x-M	M	997E
B7-4	997-M2x-M-8	M	997E
B7-5	997-M2x-A	A	997E
B7-6	997-M2x-M	M	997E
B7-7	HPE17-M1-NUS0	N/A	HPE
B7-8	HPE30-M1-NUS0	N/A	HPE
B7-9	997E17-M2x-A	A	997E
B7-10	997E30-M2x-NUS0	N/A	997E
B8-1	998-M1x-A	A	998E and 998ADE
B8-2	998-M1x-B	B	998E and 998ADE
B8-3	998-M1x-NUS0	N/A	998E and 998ADE
B8-4	998-M2x-A	A	998E and 998ADE
B8-5	998-M2x-M	M	998E and 998ADE
B8-6	998-M2x-B	B	998E and 998ADE
B8-7	998-M2x-NUS0	N/A	998E and 998ADE
B8-8	998E17-M2x-NUS0	N/A	998E
B8-9	998E17-M2x-NUS0-M	N/A	998E
B8-10	998ADE17-M2x-NUS0-M	N/A	998ADE
B8-11	998ADE17-M2x-A	A	998ADE
B8-12	998ADE17-M2x-B	B	998ADE
B8-13	998E30-M2x-NUS0	N/A	998E
B8-14	998E30-M2x-NUS0-M	N/A	998E
B8-15	998ADE30-M2x-NUS0-M	N/A	998ADE
B8-16	998ADE30-M2x-NUS0-A	N/A	998ADE
B8-17	998ADE17-M2x-M	M	998ADE

Short Name	Limit PSD Mask (Long Name)	US0 Type A/B/M	Compatible Band Plan Type

NOTE

US0 types are as follows:

- US0 type A corresponds to G.992.5 Annex A.
- US0 type B corresponds to G.992.5 Annex B.
- US0 type M corresponds to G.992.3/G.992.5 Annex M.
- US0 type N/A designates a band plan variant that does not use US0.

Fields listed in Table 7-8 are as follows:

- **Short Name**: short name of a limit PSD mask, similar to an index. "B8" refers to band plan 998. Similarly, B7 in Annex B refers to band plan 997. Carriers generally use a short name to identify a limit PSD mask. The breakpoint and breakpoint PSD value for each type of limit PSD mask are determined based on the transmission direction, upstream (VTU-R TX direction) or downstream (VTU-O TX direction), as shown in Table 7-9 and Table 7-10, respectively.

NOTE

- Due to space limitations, only parts of contents are listed in Table 7-9 and Table 7-10. For complete contents, see ITU-T Recommendation G.993.2.
- In G.993.2 Annex B type, downstream and upstream limit PSD masks use the same short names. The only difference between them lies in breakpoints. In G.993.2 Annex A type, downstream and upstream limit PSD masks use different short names. For example, the short names of downstream limit PSD masks are D-32 and D-64 and of upstream limit PSD masks are EU-32 and ADLU-32 (described in field **US0 type** in the following section).

Table 7-9 VTU-R limit PSD masks for band plan 998 and its extensions

Name	B8-1	...	B8-8	...	B8-11	...
Long Name	998-M1x-A	...	998E17-M2x-NUS0	...	998ADE17-M2x-A	...
kHz	dBm/Hz	...	dBm/Hz	...	dBm/Hz	...
0	-97.5	...	-100	...	-97.5	...
4	-97.5	...	-100	...	-97.5	...
4	-92.5	...	-100	...	-92.5	...
25.875	-34.5	...	-100	...	-34.5	...
50	-34.5	...	-100	...	-34.5	...
80	-34.5	...	-100	...	-34.5	...
120	-34.5	...	-100	...	-34.5	...
138	-34.5	...	-100	...	-34.5	...
...	...	...	...	...	...	...
24890	-100	...	-100	...	-100	...

Name	B8-1	...	B8-8	...	B8-11	...
Long Name	998-M1x-A	...	998E17-M2x-NUS0	...	998ADE17-M2x-A	...
kHz	dBm/Hz	...	dBm/Hz	...	dBm/Hz	...
25065	-100	...	-100	...	-100	...
30000	-100	...	-100	...	-100	...
30000	-110	...	-110	...	-110	...
30175	-110	...	-110	...	-110	...
≥ 30175	-110	...	-110	...	-110	...

Table 7-10 VTU-O limit PSD masks for band plan 998 and its extensions

Name	B8-1	...	B8-8	...	B8-11	...
Long Name	998-M1x-A	...	998E17-M2x-NUS0	...	998ADE17-M2x-A	...
kHz	dBm/Hz	...	dBm/Hz	...	dBm/Hz	...
0	-97.5	...	-97.5	...	-97.5	...
4	-97.5	...	-97.5	...	-97.5	...
4	-92.5	...	-92.5	...	-92.5	...
80	-72.5	...	-72.5	...	-72.5	...
...	...	...	...	...	...	...
24890	-100	...	-100	...	-100	...
25065	-100	...	-100	...	-100	...
30000	-100	...	-100	...	-100	...
30000	-110	...	-110	...	-110	...
30175	-110	...	-110	...	-110	...
≥ 30175	-110	...	-110	...	-110	...

- **Long Name**: description of a limit PSD mask. **998** and **998E17** are band plans. **NUS0** indicates that US0 is disabled. Query Table 7-9 and Table 7-10 for details of long names when planning limit PSD masks.
- **US0 Type**: specifies the US0 spectrum type for each type of limit PSD mask, described in "NOTE" of Table 7-8.
  - **A**: indicates that the US0 spectrum range is the same as that of G.992.5 Annex A, ranging from 25 kHz to 138 kHz.

  - **B**: indicates that the US0 spectrum range is the same as that of G.992.5 Annex B, ranging from 120 kHz to 276 kHz.
  - **M**: indicates that the US0 spectrum range is the same as that of G992.3/G.992.5 Annex M, ranging from 25 kHz to 276 kHz.
  - **N/A**: indicates that US0 is not enabled.
- **Compatible Band Plan Type**: indicates available global vectoring band plans based on limit PSD masks.

Activation Policies for Legacy CPEs

CPEs in a vectoring system are classified as vectoring CPEs, vectoring friendly CPEs, and legacy CPEs. Both vectoring CPEs and vectoring friendly CPEs support vectoring process flows. Legacy CPEs do not support vectoring process flows and can be activated only in G.993.2 mode.

If a legacy CPE in a vectoring system is activated in G.993.2 mode:

- Vectoring cannot eliminate the crosstalk from other CPEs to this legacy CPE and therefore cannot improve the performance of this CPE.
- Vectoring cannot eliminate the crosstalk from this legacy CPE to vectoring or vectoring friendly CPEs.

To minimize the degradation on vectoring performance caused by legacy CPEs, configure the activation policies for legacy CPEs. To do so, run the **xdsl vectoring legacy-cpe activate-policy** command.

Table 7-11 Activation policies for legacy CPEs

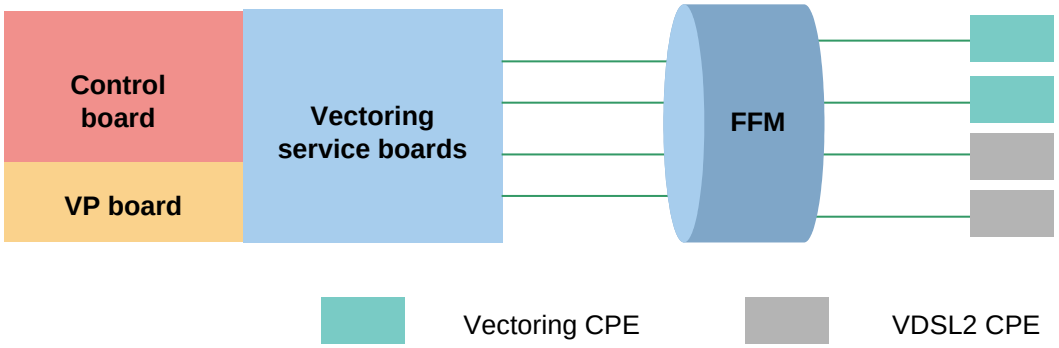
Activation policy	Description	Usage Scenario
No-limit	Allows a legacy CPE to be activated in common VDSL2 mode. In this mode, the vectoring performance is degraded.	This activation policy is used at the initial vectoring application phase. During this phase, a large number of CPEs need to be upgraded or replaced, and the vectoring performance is not of prime concern.
Limit	Allows a legacy CPE to be activated in G.993.2 mode. In addition, the AN will automatically shape the PSD of the line connected to this legacy CPE so that the CPE is activated at a low VDSL2 rate. This prevents this line from decreasing the rates of other lines. **NOTE** Legacy CPEs are activated using the limit policy by default.	This activation policy is used at the medium vectoring application phase. During this phase, only some CPEs have not been upgraded or replaced; therefore, they can be activated using this policy to adapt to the entire vectoring system.
Force-friendly-ds-limit-us	Allows a legacy CPE to be activated using the force friendly	This activation policy is used at the medium vectoring application

Activation policy	Description	Usage Scenario
	policy in the downstream direction and the automatic PSD shaping in the upstream direction. **NOTE** The force friendly policy cancels crosstalk on vectoring lines brought by legacy CPEs but does not improve the performance of legacy CPEs.	phase. During this phase, only some CPEs have not been upgraded or replaced; therefore, they can be activated using this policy to adapt to the entire vectoring system.
Force-friendly-ds-no-limit-us	Allows a legacy CPE to be activated using the force friendly policy in the downstream direction and applies no limitation in the upstream direction.	This activation policy is used at the medium vectoring application phase. During this phase, only some CPEs have not been upgraded or replaced; therefore, they can be activated using this policy to adapt to the entire vectoring system.
Block	Prohibits a legacy CPE from being activated in G.993.2 mode.	This activation policy is used in the mature vectoring application stage. During this stage, the vectoring performance is concerned and unnecessary crosstalk is better to be masked.

FFM

The force friendly mode (FFM) is a Huawei-developed policy for activating legacy CPEs. In this mode, legacy CPEs are forced to be vectoring friendly in the downstream direction, and the AN connected to these legacy CPEs automatically shapes line PSDs for these legacy CPEs, thereby minimizing vectoring performance degradation caused by these legacy CPEs.

Figure 7-23 FFM applications



In the preceding figure,

- Vectoring service boards support FFM.

- Crosstalk on vectoring lines is canceled using an algorithm. This increases vectoring line rates on the basis of without decreasing VDSL2 line rates.
  - FFM applies to the downstream direction. Specifically, the AN uses a dedicated algorithm to obtain the global crosstalk based on the crosstalk returned from certain vectoring CPEs for crosstalk cancellation.
  - FFM cannot take effect in the upstream direction because legacy CPEs cannot send their crosstalk in the upstream direction. PSD shaping is used in the upstream direction, which uses a smart algorithm to limit the frequency band generating the largest crosstalk. This minimizes vectoring performance degradation by decreasing VDSL2 line rates to a small extent.

NOTE

- A PSD is a differential of the TX power on frequencies, representing the power of a frequency, in the unit of dBm/Hz. Oppositely, the cumulative PSDs on frequencies in a spectrum band are the TX power of this spectrum band. The purpose of PSD control is to eliminate external noises and minimize crosstalk outputs. For more information about PSD, see 6.3.6 PSD Profiles.
- PSD shaping is implemented using a management information base (MIB) PSD mask. For details, see 6.3.9 MIB PSD Mask.

Fast Port Activation

Compared with the time required for activating a VDSL2 port, the time required for activating a vectoring port is longer. The reason is that both the ports newly added to a vectored group and the online ports in this group must update crosstalk, which is performed at training phase according to ITU-T Recommendation G.993.5.

The increase for the number of lines enlarges calculation volume and prolongs calculation time. In addition, the pilot sequence (PS) becomes longer accordingly. When a port joins in a vectored group, it requires obtaining error samples (ESs) multiple times during the training. Therefore, the update of line crosstalk during the training significantly prolongs the time required for a port to go online, especially when the number of online ports is large. In this case, to improve user experience, the time required for a port to go online must be shortened using a fast port activation policy.

Vectoring ports use different fast activation policies, depending on usage scenarios. To configure a fast port activation policy, run the **xdsl vectoring fast-join** command.

Table 7-12 Fast port activation policies

Policy	Description	Remarks
at-init	Vectoring gain is obtained during activation.	This is the default fast port activation policy. When this policy is used, the vectoring port is activated using the common activation process but not activated fast.
trigger	Condition-triggered policy, which includes: • **after-board-reset**: When ports on a board are activated after the board resets, the vectoring gain of these ports is obtained after these ports are activated. This speeds up port activation and	None

Policy	Description	Remarks
	shortens service recovery time after the board or system resets. • **during-bulk-init**: If the number of ports that go online in the same batch exceeds the preset threshold, the vectoring gain is obtained after these ports are activated.	
during-showtime	Vectoring gain is obtained after ports are activated and adjusted in tracking. This speeds up port activation.	None
history-coefficient	A historical coefficient is used for vectoring calculation.	• The vectoring calculation using a historical coefficient can shorten the port activation time. If the historical coefficient of showtime ports for join-in ports is unavailable for the first calculation, the join-in port performance cannot reach the optimal level immediately after going online. In this situation, tracking must be performed to improve the performance. • If the historical coefficient of the showtime ports for the join-in ports is available, the join-in port performance approaches the optimal level immediately after being activated.
fdps	Frequency dependent pilot sequence, which shortens the PS length and reduces the number of sampling points used for calculating a crosstalk cancellation coefficient. After FDPS is enabled, the coefficient precision reduces but port activation time shortens.	For details about FDPS, see PS.

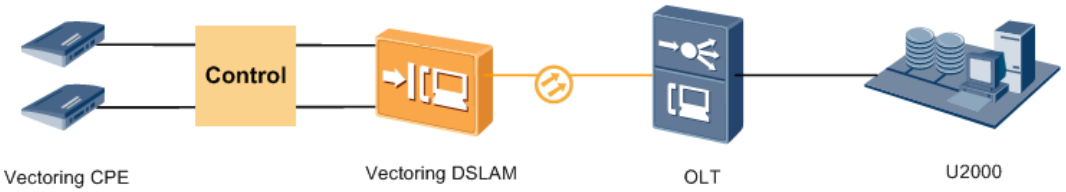
NOTE

Fast activation options **at-init**, **trigger**, and **during-showtime** cannot be selected at the same time. However, each of them can be used together with **history-coefficient** and **fdps**.

Control Policies for Frequent Online and Offline Ports

The vectoring algorithm is executed each time a port in a vectored group goes online or offline. To prevent the long-term heavy vectoring task, caused by frequent port online and offline, from affecting system functions, configure a control policy for ports that frequently go online and offline. To do so, run the **xdsl frequent-retrain-control** command.

Figure 7-24 Vectoring networking



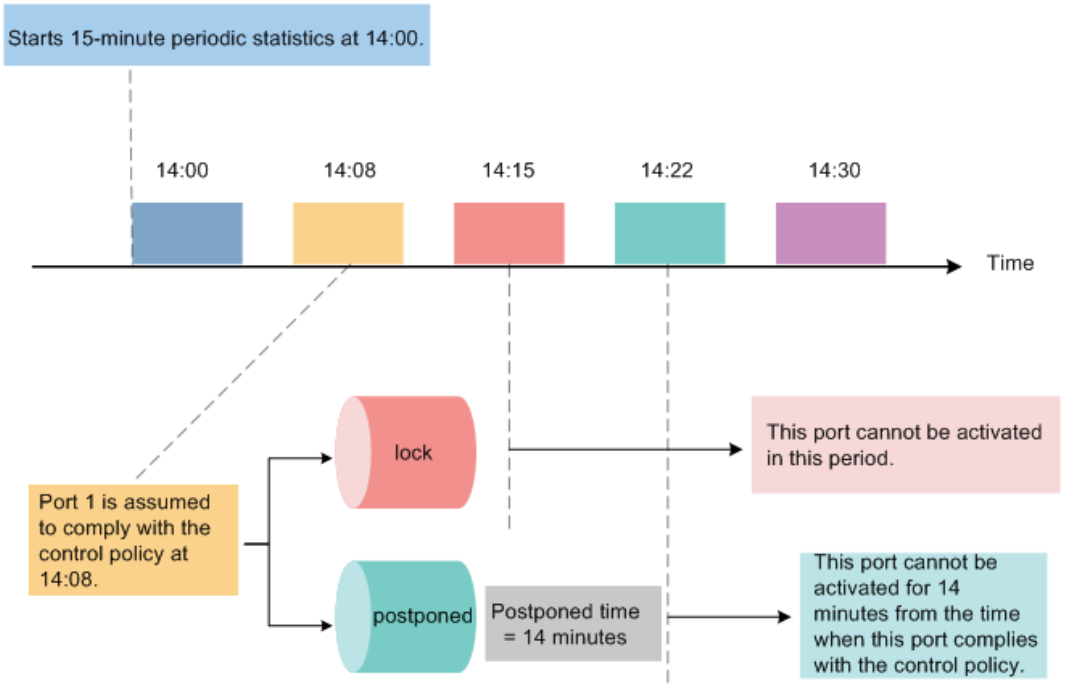
The number of frequency online and offline times of a port is assumed to reach the preset threshold N in a 15-minute statistical period. Then, a control policy can be applied on this port to improve the overall performance and stability of the vectoring system.

Table 7-13 Control policies for frequency online and offline ports

Policy	Description
no-limit	A port can be activated regardless of whether the online and offline times of this port reached the preset threshold. **NOTE** The default control policy is **no-limit**.
postponed	Within a 15-minute statistical period, if the number of online and offline times of a port reaches the preset threshold, this port cannot be activated within the configured postponed time.
lock	Within a 15-minute statistical period, if the number of online and offline times of a port reaches the preset threshold, this port cannot be activated in this period.
non-vectoring	If a vectoring port meets the requirements of triggering a control policy for frequent online and offline ports, the AN considers this port as a legacy port in the next initialization. (The restriction is removed after this vectoring port enters showtime phase again.) **NOTE** This configuration takes effect only for vectoring ports.
further-control-policy	The further control policy takes effect when the number of port retrain times reaches 3 within a detection period.

The following figure shows the differences between the **postponed** policy and the **lock** policy.

Figure 7-25 Differences between the **postponed** policy and the **lock** policy



PS

What Is PS

A PS is a binary sequence set by a VCE. When a VCE sends a PS to a VTU-R at initialization and showtime phases, each PS bit determines whether the VTU-R (depended based on the upstream PS) or VTU-O (depended based on the downstream PS) are modulated to all 0s or all 1s on all probe subcarriers of specified synchronization symbols.

NOTE

- Standard PS lengths are power of 2, ranging from 2 to 512.
- Showtime is the status of transmitting data over bearer channels after the VTU-O and VTU-R are initialized.

PS is as follows:

- The comparison between TX and RX PSs can be used to obtain error samples (ESs) for calculating a crosstalk cancellation coefficient. For details, see 7.6.2 Vectoring Principles.
- PS lengths are determined based on the number of ports. Huawei-implemented PSs are dynamically adjusted based on the number of ports.

Upstream PS Modulation

A VTU-R must be capable of modulating the upstream PS specified by a VCE on all subcarriers of upstream synchronization symbols at initialization phase or on probe subcarriers at showtime phase. An upstream PS is defined by device vendors. It is a quadrature sequence with a length of Npilot_us bits, which is sent by the VCE to the VTU-R

using the O-SIGNATURE message at initialization phase. The upstream PS is sent at a period of Npilot_us bits. It can be changed by the VCE using a flow at showtime phase.

NOTE

The vectoring management entity (VME) in the VTU-O must use the command and response to update the upstream PS and send the updated upstream PS to the VME in the VTU-R. This command can only be initiated by the VTU-O, and the VTU-R responds to the VTU-O with ACK or NACK. All commands and responses for updating upstream PSs are EOC messages.

Downstream PS Modulation

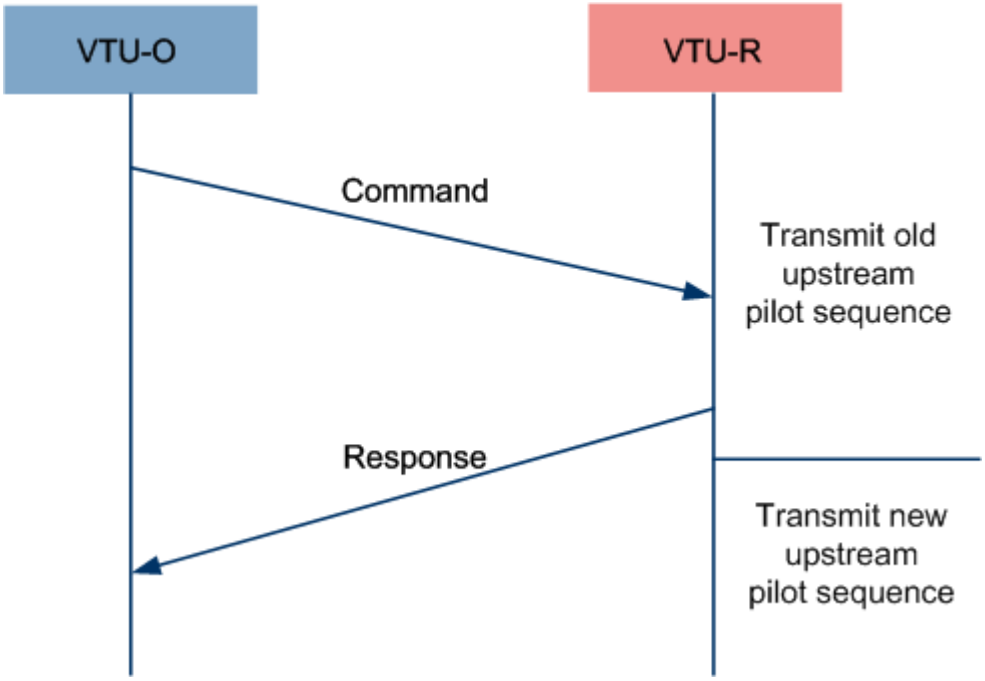
A VTU-O must be capable of modulating the downstream PS specified by a VCE on all probe subcarriers of downstream synchronization symbols at both initialization and showtime phases. A downstream PS, a binary quadrature sequence with a length of Npilot_ds bits, is determined by the VCE. The downstream PS is repeated at a period of Npilot_ds bits, unless the VCE changes the downstream PS. The VCE can change the downstream PS at any time, with no need to inform the VTU-R of the downstream PS changing only if the downstream PS length retains. At initialization phase, the VTU-O can modulate the downstream PS on all marked subcarriers of downstream synchronization symbols to all-1s or to be the same as that on probe subcarriers.

NOTE

A VCE can use a PS updating message to update an upstream PS and directly update a downstream PS without using a PS updating message.

Time sequence diagram for a PS updating EOC command and response

Figure 7-26 Time sequence diagram



NOTE

Do not frequently update an upstream PS because a stable PS facilitates FEXT channel identification.

FDPS

FDPS was promoted by ASSIA in ITU-T Recommendation draft 09GS-079 in May 2009. This feature shortens PS lengths and speeds up vectoring cancellation coefficient calculation because only PS sending is required. The PS length shortening reduces the number of samples required for calculating the vectoring cancellation coefficient. Although the coefficient precision decreases, the time required for activating a port is shortened.

NOTE

- Upstream FDPS can be supported only by dedicated CPEs.
- Downstream FDPS must be supported by the access device and Huawei devices have supported this function.
- After FDPS is enabled, vectoring performance is degraded at join-in phase. Therefore, the port activation rate is lower than that before FDPS is enabled. However, the port activation rate can be rapidly increased using tracking.

FDPS is one the methods for fast activating vectoring. For details about other fast vectoring activation configurations, see Fast Port Activation.

Crosstalk Matrix

According to ITU-T Recommendation G.993.5, a channel matrix represents the FEXT from other lines in the bundle that interferes with the VDSL2 performance on a particular line on all subcarriers.

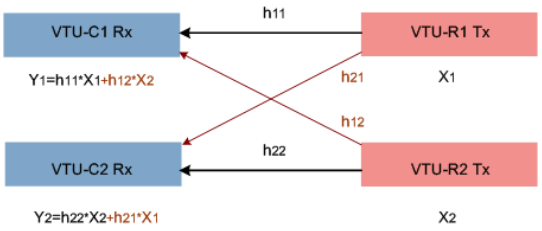
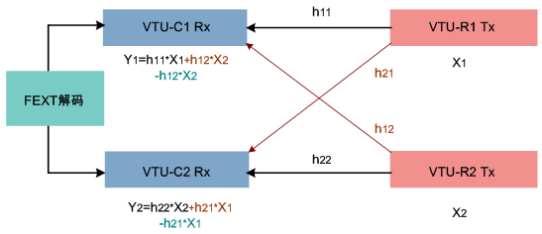
- Huawei considers a channel matrix as a crosstalk matrix, representing the crosstalk on a port caused by other ports in a vectored group.
- A crosstalk matrix is calculated at showtime phase, reflecting crosstalk strengths by PSD.
- A crosstalk cancellation matrix is obtained based on a crosstalk matrix, which can be considered as the relationships between crosstalk signals and crosstalk cancellation signals on lines.

The following section uses algorithms to describe vectoring principles, helping you to understand crosstalk matrices and crosstalk cancellation matrices.

According to communications rules, RX signal Yn = TX signal Xn x Channel transmission function Hnn. This document uses the upstream direction (from the CPE to the CO) of two DLS lines as an example for analysis.

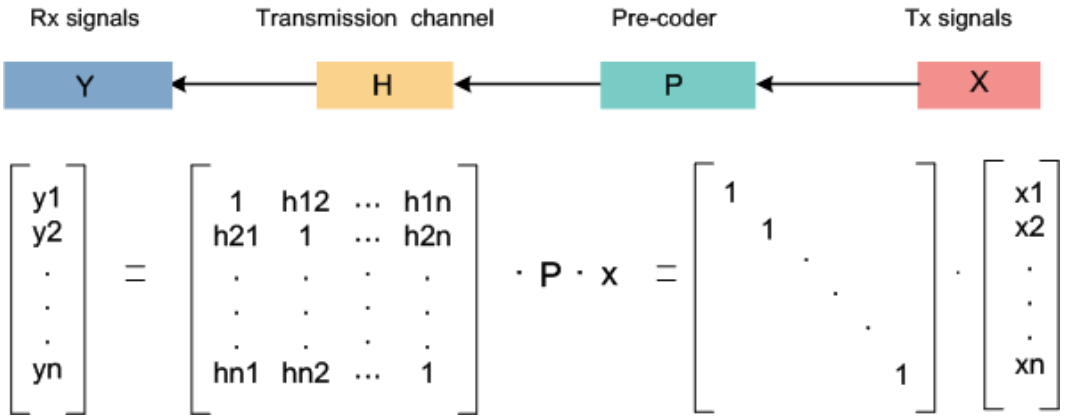
Table 7-14 Crosstalk matrix diagrams

Diagram	Description
❶ Upstream direction, without considering FEXT 	In ideal transmission without crosstalk, $Y_n = H_{nn} \times X_n$.

Diagram	Description
❷ Upstream direction, considering FEXT 	The FEXT causes distortions of (h_{12} x X_2) and (h_{21} x X_1) for Y_1 and Y_2, respectively. The matrix marked by red h_{ij} is the crosstalk matrix.
❸ Upstream direction, cancelling FEXT 	After vectoring is enabled to cancel FEXT, signal distortions are canceled and original signals are restored. The matrix marked by blue h_{ij} is the crosstalk matrix.

The rules used in the downstream direction are similar to those used in the upstream direction, as shown in the following figure.

Figure 7-27 Mathematical diagram



NOTE

- Matrix P is the crosstalk cancellation matrix.
- Theoretically, a crosstalk cancellation matrix can completely cancel a crosstalk matrix. However, this is only an expectation because it cannot be implemented due to many factors, such as poor line quality. The calculation precision for crosstalk cancellation matrices is the most difficult and major concern in applying vectoring to cancel crosstalk.

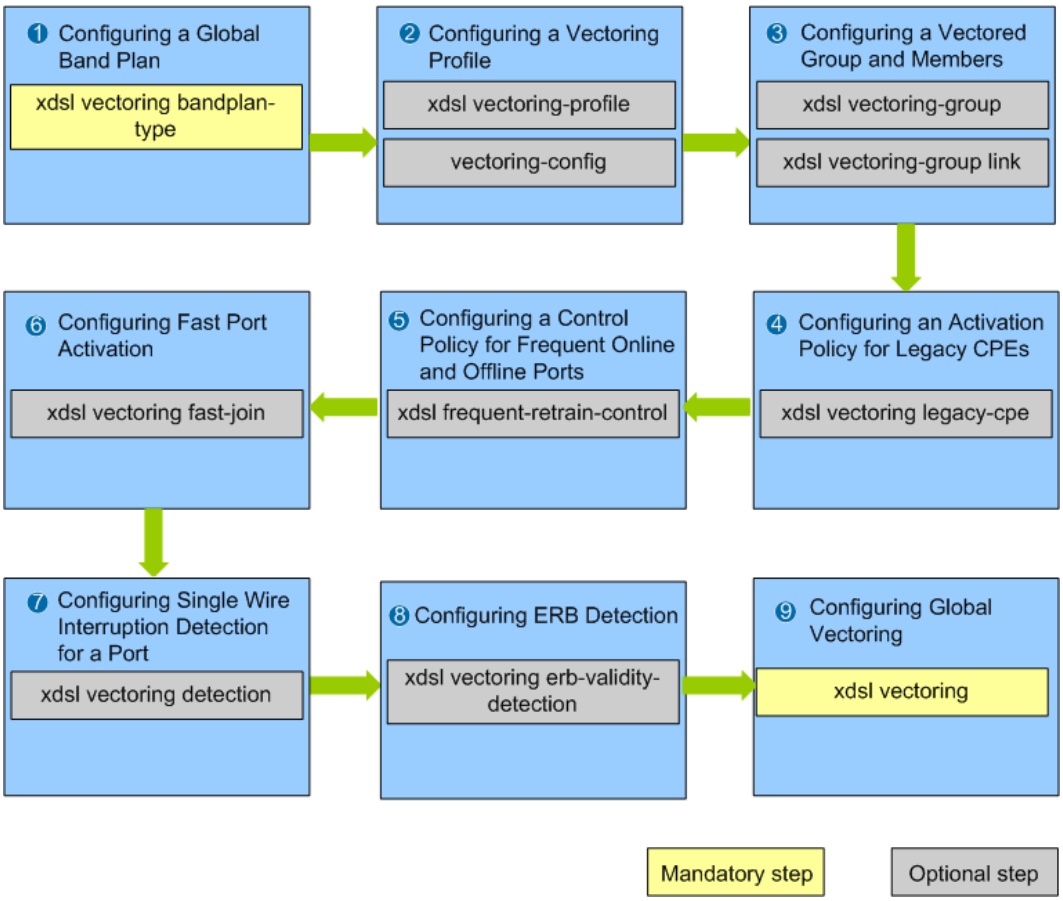
7.7 Vectoring Deployment

7.7.1 Vectoring Configuration Guide

Configuration Panorama

Figure 7-28 shows the vectoring configuration panorama.

Figure 7-28 Vectoring configuration panorama



NOTE

- Most vectoring configurations are optional. Users can configure vectoring based on actual network situations and networking planning requirements.
- The AN has default configurations for the global vectoring band plan. However, users are recommended to configure the global band plan based on the frequency band type in the VDSL2 profile bound to ports. For mapping between port profiles and global vectoring band plans, see Table 7-8.
- NLV configuration process is the same as SLV configuration process. In NLV configuration, vectoring parameters are configured on two SLV devices.NLV functions properly only if both of the following requirements are met:
- Vectoring has been enabled on both SLV devices.
- Both SLV devices have been securely connected in **错误！未找到引用源。**.

Configuring a Global Band Plan

Before enabling vectoring, configure the global band plan type. If the band plan type configured in the vectoring profile bound to a port is incompatible with the global vectoring band plan type, downstream and upstream frequency bands overlap and the port cannot be activated.

Procedure

Run the **xdsl vectoring bandplan-type** command to configure a global vectoring band plan type.

Table 7-15 Key parameter

Parameter	Description
Global vectoring band plan type	The parameter value can be 997e, 998ade, 998e, 998ade17spe-1, hpe, Annex A, or Annex C. **NOTE** • Default value: 998ade. • 998ade17spe-1: indicates the type of the vectoring global band plan, which is defined by Huawei.
US0 type	The parameter value can be type-a, type-b, type-m, or type-n/a. **NOTE** • Default value: type-a. • When the global vectoring band plan type is 997e, 998ade, hpe, or 998e, configure this parameter.

Step 2 Run the **display xdsl vectoring config** command to query the configured global vectoring band plan type.

NOTE
This command can be executed in three VDSL2 management modes: TR129, TR165, and TI. Therefore, identify the VDSL2 management mode before running this command. You can run the **displayvdsl mode** command to query the current VDSL2 management mode.

Step 3 Run the **display xdsl vectoring line-info** command to check whether the band plan type in the vectoring profile bound to the port is compatible with the global vectoring band plan type.

----End

Example

The following is an example of the configurations used to configure a global band plan:

- Band plan type: 998ade
- US0 type: type A

```
huawei(config)#xdsl vectoring bandplan-type 998ade us0-type type-a
huawei(config)#display xdsl vectoring config
 ------------------------------------------------------------------------------
  ......
  Vectoring bandplan-type            : 998ADE
  Vectoring US0 type                 : US0-type A
```

```
 ......
 --------------------------------------------------------------------------------
huawei(config)#display xdsl vectoring line-info all
 --------------------------------------------------------------------------------
  F/ S/ P    Vectoring    Vectoring    Vectoring    BandPlan    CPE type
             Profile      Group ID     Port-index   Compatible
 --------------------------------------------------------------------------------
  0/ 4/ 0          1            1            0           Y  vectoring
 --------------------------------------------------------------------------------
```

Exception Handling

Symptom	Cause	Handling Method
A port fails to activate.	The band plan type in the vectoring profile bound to the port is incompatible with the global vectoring band plan type. The **display xdsl vectoring line-info all** command output shows that **BandPlan Compatible** is **N** for this port.	1. Run the **display port state** and **display vdsl line-template** commands to query the vectoring profile used to activate the port. 2. Run the **display vdsl line-profile** command to query the band plan type configured in the vectoring profile bound to the port. 3. Query Table 7-8 to obtain the correct global vectoring band plan type. 4. Run the **xdsl vectoring bandplan-type** command to configure a correct band plan type. 5. Run the **display xdsl vectoring line-info all** command to verify that **BandPlan Compatible** is **Y** for this port.

Configuring a Vectoring Profile

To configure a vectoring profile, setting of parameters including downstream and upstream crosstalk cancellation status and a vectoring-legacy CPE activation policy is required.

Procedure

Run the **xdsl vectoring-profile add** command to create a vectoring profile and configure required parameters.

The AN uses vectoring profile 1 by default. Add other profiles if the default profile cannot meet requirements.

Table 7-16 Key parameter

Parameter	Description
control	• Enables or disables downstream crosstalk cancellation. It determines whether downstream crosstalk cancellation calculation results are applied on a port in the vectored group.

Parameter	Description
	The corresponding parameter is **fext-cancel-control-ds**. • Enables or disables upstream crosstalk cancellation. It determines whether upstream crosstalk cancellation calculation results are applied on a port in the vectored group. The corresponding parameter is **fext-cancel-control-us**.
activate-policy	Determines a vectoring-legacy CPE activation policy. **NOTE** For details, see .
vectoring-mode	Determines the vectoring mode. • Vectoring • Full-friendly • Friendly-ds
cable-type	Determines the cable type. It can be atis, quad, paper-insulated, tp100, or other.
legacy-ratio	Determines the ratio of legacy CPEs.

NOTE
After the vectoring profile is created, you can run the **xdsl vectoring-profile modify** command to modify profile configurations or run the **xdsl vectoring-profile delete** command to delete profile configurations.

Step 2 Run the **vectoring-config** command to bind the vectoring profile to a VDSL2 port.

After the binding, the profile parameters immediately take effect on the port. Default vectoring profile 1 is bound to all VDSL2 ports by default.

Step 3 Run the **display xdsl vectoring-profile** command to query parameters configured in the vectoring profile.

----End

Example

The following is an example of the configurations used to configure a vectoring profile:

- Crosstalk cancellation in the upstream and downstream directions: enable
- Profile name: huawei

```
huawei(config)#xdsl vectoring-profile add control disable enable name huawei
huawei(config-if-vdsl-0/4)#vectoring-config all profile-index 1
huawei(config)#display xdsl vectoring-profile 1
```

Exception Handling

Symptom	Cause	Handling Method
Failed to bind a vectoring profile to a	The vectoring profile is bound to a VDSL2 port not	1. Run the **xdsl vectoring-group link add** command to add the VDSL2 port to the vectored group.

Symptom	Cause	Handling Method
VDSL2 port.	in the vectored group.	2. Run the **vectoring-config** command to bind the vectoring profile to this VDSL2 port.

Configuring a Vectored Group and Members

After creating a vectored group, add ports to this vectored group. The AN collects crosstalk information about member ports, performs vectoring calculation, and cancels crosstalk on member ports.

Procedure

Run the **xdsl vectoring-group add** command to create a vectored group and configure required parameters.

The AN creates vectored group 1 by default during the initialization.

Table 7-17 Key parameter

Parameter	Description
not-required-bands-ds	Specifies the frequency band that does not require the crosstalk cancellation in the downstream direction.
not-required-bands-us	Specifies the frequency band that does not require the crosstalk cancellation in the upstream direction.
protection-of-vectoring-lines	Enables or disables the upstream and downstream vectoring line protection. If a line bundle consists of vectoring lines and ADSL, ADSL2, or ADSL2+ lines: • Specifying the frequency band that does not require the crosstalk cancellation in the downstream or upstream direction and enabling the vectoring line protection can ensure the stability of vectoring lines. • Specifying the frequency band that does not require the crosstalk cancellation in the downstream or upstream direction but disabling the vectoring line protection affect the stability of vectoring lines and even lead to error codes and user offline.

NOTE
- The AN supports only vectored group 1. All cables are added to vectored group 1 by default. The AN allows configurations of other vectored groups. However, the configurations do not take effect.
- You can run the **xdsl vectoring-group modify** command to modify vectored group 1.
- The default vectored group 1 cannot be deleted.

Step 2 Run the **xdsl vectoring-group link add** command to add vectored group members.

The AN automatically binds vectoring profile 1 to the newly added members.

Step 3 Run the **display xdsl vectoring-group** command to query information about the vectored group.

----End

Example

The following is an example of the configurations used to configure a vectored group:

- Vectored group: 1
- Frequency bands that do not require the crosstalk cancellation in the downstream direction: 33, 100-700, and 1216-1961
- Frequency bands that do not require the crosstalk cancellation in the upstream direction: 890-900

```
huawei(config)#xdsl vectoring-group modify 1 not-required-bands-ds enable
33,100-700,1216-1961 not-required-bands-us enable 890-900
```

The following is an example of the configurations used to add member ports (in port lists 0/4: 0-15, 32-47) to vectored group 1:

```
huawei(config)#xdsl vectoring-group link add 1 0/4:0-15,32-47
huawei(config)#display xdsl vectoring-group 1
```

Exception Handling

Symptom	Cause	Handling Method
A port fails to add to a vectored group.	A port cannot be added to multiple vectored groups.	1. Run the **xdsl vectoring-group link delete** command to delete the port from the current vectored group. 2. Run the **xdsl vectoring-group link add** command to add the port to a specified vectored group.

Configuring an Activation Policy for Legacy CPEs

A legacy CPE refers to the VDSL2 CPE (complying with G.993.2) that does not support vectoring. When legacy CPEs in the vectoring system are activated in G.993.2 mode, crosstalk brought by the legacy CPEs cannot be canceled, deteriorating vectoring performance. Therefore, configure the legacy CPE activation policy to minimize impacts on vectoring performance.

Procedure

Run the **xdsl vectoring legacy-cpe activate-policy** command to configure a legacy CPE activation policy.

Table 7-18 Key parameters

Parameter	Description	Remarks
No-limit	Allows a legacy CPE to be activated in common VDSL2 mode. In this mode, the vectoring performance is degraded.	This activation policy is used at the initial vectoring application phase. During this phase, a large number of CPEs need to be upgraded or replaced, and the

Parameter	Description	Remarks
		vectoring performance is not of prime concern.
Limit	Allows a legacy CPE to be activated in G.993.2 mode. In addition, the AN will automatically shape the PSD of the line connected to this legacy CPE so that the CPE is activated at a low VDSL2 rate. This prevents this line from decreasing the rates of other lines.	This activation policy is used at the medium vectoring application phase. During this phase, only some CPEs have not been upgraded or replaced; therefore, they can be activated using this policy to adapt to the entire vectoring system.
Force-friendly-ds-limit-us	Allows a legacy CPE to be activated using the force friendly policy in the downstream direction and the automatic PSD shaping in the upstream direction. **NOTE** The force friendly policy cancels crosstalk on vectoring lines brought by legacy CPEs but does not improve the performance of legacy CPEs.	This activation policy is used at the medium vectoring application phase. During this phase, only some CPEs have not been upgraded or replaced; therefore, they can be activated using this policy to adapt to the entire vectoring system.
Force-friendly-ds-no-limit-us	Allows a legacy CPE to be activated using the force friendly policy in the downstream direction and applies no limitation in the upstream direction.	This activation policy is used at the medium vectoring application phase. During this phase, only some CPEs have not been upgraded or replaced; therefore, they can be activated using this policy to adapt to the entire vectoring system.
Block	Prohibits a legacy CPE from being activated in G.993.2 mode.	This activation policy is used in the mature vectoring application stage. During this stage, the vectoring performance is concerned and unnecessary crosstalk is better to be masked.
cable-type	Determines the cable type. It can be atis, quad, paper-insulated, tp100, or other.	None
legacy-ratio	Determines the ratio of legacy CPEs.	None

NOTE

- In TR129 and TR165 modes, the optional rate profile **limit-profile** is added.

- When the activation policy is **limit**, you can specify **reserved-band** (shaping is not performed on these reserved bands and the limit activation policy is not applied) or **blackout-band** (the limit activation policy applies to these ports).

Step 2 Run the **display xdsl vectoring config** command to query the configured legacy CPE activation policy.

----End

Example

The following is an example of the configurations used to configure a legacy CPE activation policy:

- CPE activation policy: limit
- Cable type: atis
- Legacy CPE ratio: 1% to 15%

```
huawei(config)#xdsl vectoring legacy-cpe activate-policy limit reserved-band enable
0-511 blackout-band enable 512-4095 cable-type atis legacy-ratio 0
huawei(config)#display xdsl vectoring config
 ------------------------------------------------------------------------------
  ......
  Vectoring legacy CPE activate-policy : Limit
  Vectoring legacy CPE reserved-band   : 0-511
  Vectoring legacy CPE blackout-band   : 512-4095
  ......
  Cable type                       : ATIS
  Legacy CPE ratio                  : 1%~15%
  ------------------------------------------------------------------------------
```

Exception Handling

Symptom	Cause	Handling Method
A legacy CPE fails to activate.	The activation policy for the legacy CPE is **block**.	1. Run the **display xdsl vectoring line-info** command to query the CPE type and confirm that **CPEtype** is **VDSL2**. 2. Run the **display xdsl vectoring config** command to query the legacy CPE activation policy and confirm that the legacy CPE activation policy is **Block**. 3. Run the **xdsl vectoring legacy-cpe activate-policy** command to modify the legacy CPE activation policy to a planned one.
	A legacy CPE does not support retransmission defined in its activation profile.	1. Run the **display xdsl vectoring line-info** command to query the CPE type and confirm that **CPEtype** is **VDSL2**. 2. Run the **display port state** and **display vdsl line-template** commands to query the activation profile of the port. 3. Run the **display vdsl line-profile**

Symptom	Cause	Handling Method
		command to check whether the retransmission policy is configured in the activation profile. The command output displays **G.998.4retransmission control in downstream/upstream : RTX_PREFERRED**. 4. Run the **vdsl line-profile modify** command to change the retransmission mode to **RTX_FORBIDDEN** in the VDSL2 profile.

Configuring a Control Policy for Frequent Online and Offline Ports

The vectoring algorithm is executed each time a port in a vectored group goes online or offline. To prevent the long-term heavy vectoring task, caused by frequent port online and offline, from affecting system functions, configure a control policy for ports that frequently go online and offline.

Procedure

Run the **xdsl frequent-retrain-control** command to configure a policy for controlling ports that frequently go online and offline.

This command takes effect immediately after being executed. The default control policy is **no-limit**.

Table 7-19 Key parameter

Parameter	Description
no-limit	A port can be activated regardless of whether the online and offline times of this port reached the preset threshold.
postponed	Within a 15-minute statistical period, if the number of online and offline times of a port reaches the preset threshold, this port cannot be activated within the configured postponed time.
lock	Within a 15-minute statistical period, if the number of online and offline times of a port reaches the preset threshold, this port cannot be activated in this period.
non-vectoring	If a vectoring port meets the requirements of triggering a control policy for frequent online and offline ports, the AN considers this port as a legacy port in the next initialization. (The restriction is removed after this vectoring port enters showtime phase again.) **NOTE** This configuration takes effect only for vectoring ports.
further-control-policy	The further control policy takes effect when the number of port retrain times reaches 3 within a detection period.

NOTE
The number of times a port goes online and offline excludes port activation and deactivation triggered by command execution.

Step 2 Run the **display xdsl frequent-retrain-control** command to query the policy for controlling ports that frequently go online and offline.

----End

Example

The following is an example of the configurations used to configure a control policy for ports that frequently go online and offline:

- Vectoring port: 0/4/0
- Maximum number of online and offline times within 15 minutes: 10
- Control policy: postponed
- Postponed duration: 10 minutes

```
huawei(config)# xdsl frequent-retrain-control 0/4/0 control-policy postponed 10 10
huawei(config)#display xdsl frequent-retrain-control 0/4/0
```

Exception Handling

Symptom	Cause	Handling Method
A vectoring port fails to activate.	When the control policy for the vectoring port is set to lock or postpone, if the number of online and offline times of this port reaches the preset threshold due to unstable line conditions, the port is locked in the preset period and cannot be activated.	1. Run the **display xdsl frequent-retrain-control** command to check whether **ControlState** is **Y**. 2. If **ControlState** is **Y**, the port cannot be activated in the preset period of the controlling state. The port can be activated only after the preset period of the controlling state elapses.

Configuring Fast Port Activation

To obtaining vectoring gains at different phases, configure fast activation options for the port.

Procedure

Run the **xdsl vectoring fast-join** command to configure fast activation options for a vectoring port.

Table 7-20 Key parameter

Parameter	Description	Remarks
at-init	Vectoring gain is obtained	This is the default fast port

Parameter	Description	Remarks
	during activation.	activation policy. When this policy is used, the vectoring port is activated using the common activation process but not activated fast.
trigger	Condition-triggered policy, which includes: • **after-board-reset**: When ports on a board are activated after the board resets, the vectoring gain of these ports is obtained after these ports are activated. This speeds up port activation and shortens service recovery time after the board or system resets. • **during-bulk-init**: If the number of ports that go online in the same batch exceeds the preset threshold, the vectoring gain is obtained after these ports are activated.	None
during-showtime	Vectoring gain is obtained after ports are activated and adjusted in tracking. This speeds up port activation.	None
history-coefficient	A historical coefficient is used for vectoring calculation.	• The vectoring calculation using a historical coefficient can shorten the port activation time. If the historical coefficient of showtime ports for join-in ports is unavailable for the first calculation, the join-in port performance cannot reach the optimal level immediately after going online. In this situation, tracking must be performed to improve the performance. • If the historical coefficient of the showtime ports for the join-in ports is available, the join-in port performance approaches the optimal level immediately after being activated.
fdps	Frequency dependent pilot	For details about FDPS, see PS.

Parameter	Description	Remarks
	sequence, which shortens the PS length and reduces the number of sampling points used for calculating a crosstalk cancellation coefficient. After FDPS is enabled, the coefficient precision reduces but port activation time shortens.	

NOTE

Fast activation options **at-init**, **trigger**, and **during-showtime** cannot be selected at the same time. However, each of them can be used together with **history-coefficient** and **fdps**.

Step 2 Run the **display xdsl vectoring fast-join config** command to query fast activation options for a vectoring port.

----End

Example

The following is an example of the configurations used to configure the fast activation for a vectoring port:

```
huawei(config)#xdsl vectoring fast-join gain-phase trigger after-board-reset enable
during-bulk-init 64 history-coefficient
both join-gap-wait-time 5 join-max-wait-time 15 tracking-period 30
huawei(config)#display xdsl vectoring fast-join config
```

Exception Handling

Symptom	Cause	Handling Method
A newly added port fails to activate.	**join-max-wait-time** is the maximum duration in which a port waits for other ports in the same join group. If the parameter value is excessively large (for example, 60s), the OPV-1 phase times out before all data is processed. As a result, follow-up phases cannot be performed and the port cannot be activated.	1. Run the **display xdsl vectoring fast-join config** command to check the value of join-max-wait-time (**Vectoring join in max wait time** in the command output). 2. Run the **xdsl vectoring fast-join** command to change the value of join-max-wait-time (the default 15s is recommended).
A vectoring port fails to activate the keeps in the activating state.	After the fast activation is enabled, the vectoring port fails to active if the historical coefficient is incorrect.	1. Run the **display xdsl vectoring fast-join config** command to check whether the fast activation function is enabled (**Vectoring gain phase** in the command output). 2. Run the **xdsl vectoring fast-join**

Symptom	Cause	Handling Method
		gain-phase at-init command to disable the fast activation function. Before the disabling, communicate with customers about whether they are sure to disable the fast activation and inform them of impacts after the fast activation is disabled.

Configuring Single Wire Interruption Detection for a Port

To prevent a vectoring port with a single interrupted wire from affecting services carried by other vectoring ports, configure single wire interruption detection for the vectoring port.

Procedure

Run the **xdsl vectoring detection** command to configure the single wire interruption detection for a vectoring port.

Table 7-21 Key parameter

Parameter	Description	Remarks
single-wire-interruption	Indicates single wire interruption detection. This function is disabled by default.	After the single wire interruption detection and MELT test functions are enabled for a port, the AN performs the following operations: • After determining that a single wire is interrupted, the AN reports a vectoring single wire fault alarm, sets the port to work in legacy CPE mode, and allows the port to be activated with a certain frequency spectrum. • After determining that the single wire fault does not occur, the AN performs a MELT test on the single wire. If the test result shows that the single wire is faulty, the AN does not perform any operations. If the test result shows that the single wire is functional, the AN reports a single wire fault clear alarm and removes restrictions on the port.
melt-detection-switch	Indicates MELT detection. This function is disabled by default.	

NOTE

- The enabled single wire interruption detection takes effect only after vectoring is enabled.
- The MELT test takes a long time. Perform a MELT test based on actual requirements.

Step 2 Run the **display xdsl vectoring detection config** command to query configurations about the single wire interruption detection for a vectoring port.

----End

Example

The following is an example of the configurations used to enable the vectoring single wire interruption detection:

```
huawei(config)#xdsl vectoring detection single-wire-interruption enable melt enable
huawei(config)#display xdsl vectoring detection config
  ------------------------------------------------------------------
  XDSL vectoring single-wire-interruption detection   : Enable
  Melt detection                                      : Enable
  ------------------------------------------------------------------
```

Configuring ERB Detection

After you configure the REIN ERB detection and the ERB packet drop thresholds for the join-in and tracking flows, the vectoring performance can be enhanced.

Procedure

Run the **xdsl vectoring erb-validity-detection** command to enable ERB detection.

Table 7-22 Key parameters

Parameter	Description
level	ERB detection for the REIN protection. ERB detection is disabled by default. • 0: disables ERB detection. • 1: enables ERB detection. **NOTE** After ERB detection is enabled, the system detects ERB packets. Specifically, the system checks validity of headers and the content format of ERB packets.
erb-drop-threshold	Threshold of the ERB packet drop rate. • Threshold 75% is recommended for the join-in flow. • Threshold 7% is recommended for the tracking flow. **NOTE** If the ratio of dropped ERB packets on a port exceeds the preset threshold, the system does not calculate and update the crosstalk cancellation coefficient for this port.

Step 2 Run the **display xdsl vectoring erb-validity-detection** command to check whether ERB detection is enabled.

----End

Example

The following is an example of the configurations used to configure the ERB detection function:

- ERB detection level: 1
- ERB packet drop rate threshold in the join-in flow: 76
- ERB packet drop rate threshold in the tracking flow: 8

```
huawei(config)#xdsl vectoring erb-validity-detection 1 erb-drop-threshold 76 8
huawei(config)#display xdsl vectoring erb-validity-detection
  ---------------------------------------------------------------------------
  Detection Level                  : 1
  ERB Drop Threshold During Joining    : 76
  ERB Drop Threshold During Tracking   : 8
  ---------------------------------------------------------------------------
```

Exception Handling

Symptom	Cause	Handling Method
A vectoring port cannot be properly activated after REIN noises are added to the vectoring port.	ERB detection is disabled on the vectoring port.	1. Run the **display xdsl vectoring erb-validity-detection** command to check whether ERB detection is enabled. The command output shows that **DetectionLevel** is **0**, indicating that ERB detection is disabled. 2. Run the **xdsl vectoring erb-validity-detection** command to enable ERB detection.
	The ERB drop rate threshold is inappropriate.	1. Run the **display xdsl vectoring erb-validity-detection** command to query the ERB drop rate threshold. 2. Run the **xdsl vectoring erb-validity-detection** command to modify the ERB drop rate threshold. **NOTE** Values in Table 7-22 are recommended. If strong crosstalk exists, reduce the ERB packet drop rate threshold in the tracking flow.

Configuring Global Vectoring

To use vectoring to process signals in both downstream and upstream directions to cancel FEXT and improve VDSL2 line performance, configure the global vectoring.

Procedure

Run the **xdsl vectoring** command to enable global vectoring.

NOTE
Enable vectoring after configuring the vectoring profile, vectored group, and global band plan.

Step 1 Run the **display xdsl vectoring config** command to check whether global vectoring is enabled.

----End

Example

The following is an example of the configurations used to enable global vectoring:

```
huawei(config)#xdsl vectoring enable
  Warning: this operation will reactivate all the VDSL ports supporting
vectoring and may take several minutes.
  Are you sure to continue? (y/n)[n]:y

huawei(config)#display xdsl vectoring config
 --------------------------------------------------------------------------------
  Global vectoring configuration      : Enable
  ......
  --------------------------------------------------------------------------------
```

Exception Handling

Symptom	Cause	Handling Method
Vectoring fails to enable, and lines are still working in VDSL2 mode.	The band plan type in the vectoring profile bound to the port is incompatible with the global vectoring band plan type. The **display xdsl vectoring line-info all** command output shows that **BandPlan Compatible** is **N** for this port.	1. Run the **display port state** and **display vdsl line-template** commands to query the vectoring profile used to activate the port. 2. Run the **display vdsl line-profile** command to query the band plan type configured in the vectoring profile bound to the port. 3. Check Table 7-8 to obtain the correct global vectoring band plan type. 4. Run the **xdsl vectoring bandplan-type** command to configure a correct band plan type. 5. Run the **display xdsl vectoring line-info all** command to verify that **BandPlan Compatible** is **Y** for this port. 6. Run the **xdsl vectoring enable** command to enable global vectoring.

Querying Vectoring Configurations

After vectoring is enabled, you can run commands to query crosstalk coefficient, crosstalk strength, whether the crosstalk is canceled, interface and synchronization status between the vectoring processing board and vectoring service boards.

Procedure

Run the **display xdsl subcarrier** { **xlin-ds** | **xlin-us** } command in privilege mode or run the **display xdsl subcarrier** { **xlog-ds** | **xlog-us** } command in diagnose mode to query impact coefficient estimated by the vectoring system of other lines to a target line, that is, the crosstalk coefficient.

NOTE

This command displays Xlin information and Xlog information in different formats about crosstalk impact between lines.

Table 7-23 Xlin and Xlog information

Item	Query Result
Xlin information (both downstream and upstream information is displayed, the downstream information is used as an example) **NOTE** Xlin information is displayed in a(n) + j*b(n) format.	---------- ---------- Downstream FEXT coupling represented as ((XLINSCds/2^15)*((a(n)+j*b(n))/2^15)) Subcarrier(n): a(n) + j*b(n) ---------- ---------- 0: -626 - j*481 1: -481 + j*517 2: 517 - j*225 3: -225 - j*18 4: -18 - j*10 5: -10 - j*6 ......
Xlog information (both downstream and upstream information is displayed, the downstream information is used as an example) **NOTE** Xlog information is displayed in format of amplitude|phase.	---------- ---------- 0: 70.32 | 180 1: 60.32 | 80 2: 50.32 | 60 3: 40.32 | 50 4: 30.32 | -80 5: 20.32 | -180 ......

Step 2 Run the **display xdsl vectoring crosstalk-coupling-matrix** command in privilege mode to query the impact of other ports on the target port.

The result is a group of crosstalk values. You can also check whether the crosstalk on the egress port has been canceled.

Table 7-24 Crosstalk matrix information

Item	Query Result
Crosstalk matrix information (both upstream and downstream information is displayed, the downstream information is used as an example) **NOTE** • **Average-Magnitude**: indicates the crosstalk strength in unit of dBm/Hz. The crosstalk strength is sequenced in descending order.	---------- ---------- The disturber line listed in descending order by influence ---------- ---------- F/S/P Average-Magnitude (dBm/Hz) Cancellation Status

Item	Query Result
That is, the AN preferentially cancels strongest crosstalk on a port. • **Cancellation status**: indicates the crosstalk cancellation status. **Y** indicates that crosstalk has been canceled. **N** indicates that crosstalk is not canceled.	---------- ---------- 0/ 4/ 3 -106.0 Y 0/ 4/ 5 -110.0 Y 0/ 4/11 -110.0 Y 0/ 4/ 8 -112.0 Y 0/ 4/ 6 -114.0 Y

Step 3 Run the **display xdsl vectoring state** command in diagnose mode to query interface and synchronization status between the vectoring processing board and vectoring service boards.

Table 7-25 Interface and synchronization status

Item	Query Result
Interface and synchronization status **NOTE** • If the port of the vectoring processing board or VDSL2 board is in the down or asynchronous state, the vectoring function will be affected. • Knowing the vectoring port status facilitates the maintenance of the vectoring processing board and vectoring service boards and locating of vectoring-related faults.	---------- ---------- F/S VP vectoring VDSL vectoring Sync interface state interface state state ---------- ---------- 0/4 Up Up Synchronized ---------- ----------

----End

7.7.2 Vectoring Configuration Example

Data Plan

Table 7-26 Key parameters of vectoring configurations

Item	Data	Description
Global vectoring configurations	enable	None
Global band plan	• Band plan type: 998ade	None

Item	Data	Description
	• US0 type: type-a	
Policy for activating a CPE	no-limit	The policy for activating a legacy CPE is configured as **no-limit** at initialization stage of vectoring application.
Policy for controlling frequent online and offline on ports	no-limit	None
Vectoring group	Group ID: 1	All ports are added to the default vectoring group 1.
Vectoring profile	Profile ID: 1 The upstream/downstream crosstalk cancellation function in default Profile 1 are as follows: • Upstream crosstalk cancellation: enable • Downstream crosstalk cancellation: enable	None

Table 7-27 Key parameters of other configurations

Item	Data
Dialup mode for Internet access	PPPoE
Anti-theft and roaming of user account through PITP	enable
VDSL2 mode	TR129
VLAN	Service VLAN ID: 50, type: smart
Service port index	3

Configuration Example

On a fiber to the building (FTTB) or fiber to the curb (FTTC) network, a OLT has the vectoring function enabled and provides the Point-to-Point Protocol over Ethernet (PPPoE) Internet access service for VDSL2 users. This topic describes how to configure Internet access service on such a OLT.

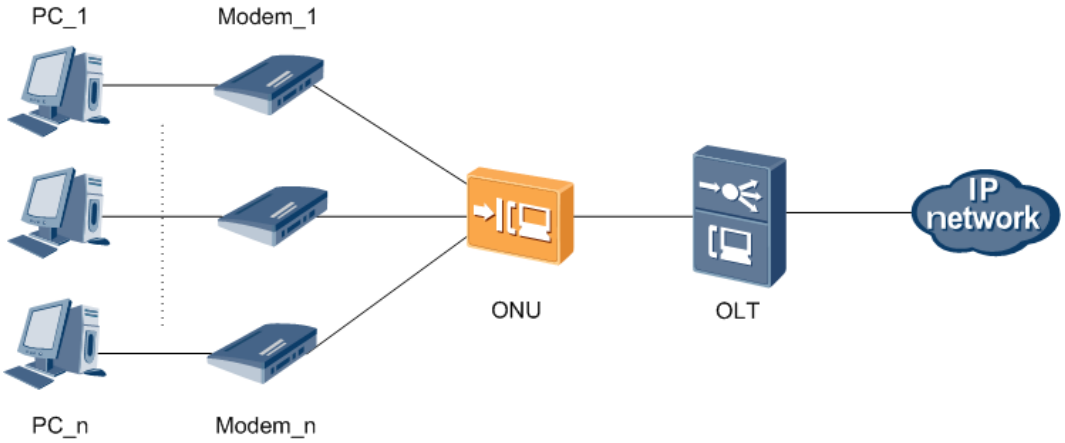
Service Requirements

- In a new VDSL2 vectoring office, all VDSL2 lines connected to the OLT are physically bundled together, and all users connect to the Internet in PPPoE mode.

- A customer premises equipment (CPE) that supports the vectoring function and a CPE that does not support the vectoring function are connected to the OLT. (A CPE that does not support the vectoring function is called a vectoring legacy CPE.)
- Different virtual local area networks (VLANs) are used to differentiate access users.
- The user access rates are not limited to prevent the vectoring performance from being affected.
- The vectoring function takes effect in upstream and downstream directions of the VDSL2 lines to cancel the far-end crosstalk (FEXT).
- User accounts must be protected against theft and roaming.
- The VDSL2 mode is set to TR129.

Figure 7-29 shows a VDSL2 Internet access service network that uses a vectoring-enabled OLT.

Figure 7-29 Internet access service network that uses a vectoring-enabled OLT



Prerequisite

The user name and password must be configured on the broadband remote access server (BRAS) for the BRAS to implement the Authentication, Authorization and Accounting (AAA) function. To implement AAA, the BRAS needs to identify the VLAN tags carried in the user packets forwarded by the OLT upstream.

Procedure

Create a service VLAN (SVLAN) and add an uplink port to the SVLAN.

Create Smart SVLAN 50 and add uplink port 0/9 to SVLAN 50.

```
huawei(config)#vlan 50 smart
huawei(config)#port vlan 50 0/9 0
```

Step 1 Configure a VDSL2 access mode.

1. Configure a VDSL2 profile. For details, see **Configuring the VDSL2 Profile**. Set the IDs of the VDSL2 line profile, VDSL2 channel profile, and VDSL2 line template to 3, channel mode to interleave, maximum downstream interleave delay to 8 ms, maximum

upstream interleave delay to 2 ms, noise margin to 6 dB, minimum downstream impulse noise protection (INP) to 4, and minimum upstream INP to 2.

```
huawei(config)#vdsl line-profile quickadd 3 snr 60 0 300 60 0 300
huawei(config)#vdsl channel-profile quickadd 3 path-mode ptm interleaved-delay 8
2 inp 4 2
huawei(config)#vdsl line-template quickadd 3 line 3 channel1 3 100 100
```

2. Activate VDSL2 port 0/4/1, and bind the configured VDSL2 line template 3 and the default VDSL2 alarm template 1 to this port.

```
huawei(config)#interface vdsl 0/4
huawei(config-if-vdsl-0/4)#deactivate 1
huawei(config-if-vdsl-0/4)#activate 1 template-index 3
huawei(config-if-vdsl-0/4)#alarm-config 1 1
huawei(config-if-vdsl-0/4)#quit
```

3. Run the **display traffic table** command to query the configured traffic profile in the system.

```
huawei(config)#<b>display traffic table ip from-index 0  </b>
{ <cr>|to-index<K> }:

 Command:
       display traffic table ip from-index 0
 ------------------------------------------------------------------------------
  TID CIR     CBS     PIR     PBS     Pri Copy-policy       Pri-Policy
     (kbps)   (bytes)  (kbps)   (bytes)
 ------------------------------------------------------------------------------
   0 512     18384    1024    36768     6 -                  tag-pri
   1 1024    34768    2048    69536     0 -                  tag-pri
   2 2048    67536    4096    135072    0 -                  tag-pri
   3 4096    133072   8192    266144    4 -                  tag-pri
   4 8192    264144   16384   528288    4 -                  tag-pri
   5 16384   526288   32768   1024000   4 -                  tag-pri
   6 off     off      off     off       0 -                  tag-pri
 ------------------------------------------------------------------------------
 Total Num : 7
```

The Internet access service requires that the user access rates not be limited. The query result shows that traffic profile 2 meets the requirements.

NOTE

- If an expected traffic profile is not available in the system, run the **traffic table** command to configure one.
- On the OLT, the user access rate can be limited by either a traffic profile or a VDSL2 line profile. When both profiles are configured, the smaller rate configured in the two profiles is used as the user bandwidth.

4. Run the **service-port** command to create a service port on user port 0/4/1. The traffic profile is profile 2 that meets the service requirements, SVLAN is 50, VDSL2 channel mode is PTM, and service port index is 3. To facilitate maintenance, the service port description information is also configured.

```
huawei(config)#service-port 3 vlan 50 vdsl mode ptm 0/4/1 multi-service user-vl
an untagged inbound traffic-table index 2 outbound traffic-table index 2
huawei(config)#service-port desc 3 description Vlanid:50/vdsl
```

Step 2 Configure a security mode for the user account.

The Policy Information Transfer Protocol (PITP) P mode can be used to protect user accounts against theft and roaming. The relay agent info option (RAIO) mode can be customized based on site requirements. This procedure uses the common mode as an example.

```
huawei(config)#pitp enable pmode
huawei(config)#raio-mode common pitp-pmode
```

NOTE

For details about the PITP configuration for user account security, see **Configuring Anti-Theft and Roaming of User Account Through PITP**.

Step 3 Configure the vectoring function.

1. Set the global bandplan to default values (998ade for bandplan type and type-a for US0 type).
2. Use the default vectoring group (group 1) to cancel the crosstalk on all frequency bands.

```
huawei(config)#display xdsl vectoring-group 1
  ------------------------------------------------------------------------
  Vectoring group index    : 1
  Lines in a vectoring group:
    0/4
  FEXT cancellation not required frequency bands downstream: 33,100-700,1216-1961
  Vectoring lines protection switch downstream          : Enable
  FEXT cancellation not required frequency bands upstream  : -
  Vectoring lines protection switch upstream            : Disable
  ------------------------------------------------------------------------
```

3. Run the **display xdsl vectoring-profile** command to query the default vectoring profile (profile 1).

```
huawei(config)#display xdsl vectoring-profile 1

------------------------------------------------------------------------------
  Profile index   : 1
  Profile name    : DEFVAL
  FEXT cancellation control upstream      : Enable
  FEXT cancellation control downstream    : Enable

------------------------------------------------------------------------------
```

   The query result shows that vectoring profile 1 meets the requirements and can be used.

4. Configure the vectoring legacy CPE activation policy to no-limit in consideration that the vectoring function is currently in the beginning phase of applications.

```
huawei(config)#xdsl vectoring legacy-cpe activate-policy no-limit
```

5. Enable the global vectoring function.

```
huawei(config)#xdsl vectoring enable
```

Step 4 Save the data.

```
huawei(config)#save
```

----End

Verification

- Setp 1: Configure the dialup user name and password on the modem. Ensure that the configurations be the same as the user name and password configured on the BRAS.

- Step 2: After the settings on the modem are completed, dialing is initialized, a network connection is no-limitmatically set up, and the user can access the Internet.
- Step 3: Log in to a network rate test website to test the rate. It is found that the upstream and downstream rates are 95% higher than the rates when the vectoring function is not enabled on the device.

Configuration File

```
vlan 50 smart
port vlan 50 0/9 0
vdsl line-profile quickadd 3 snr 60 0 300 60 0 300
vdsl channel-profile quickadd 3 path-mode ptm interleaved-delay 8 2 inp 4 2
vdsl line-template quickadd 3 line 3 channel1 3 100 100
interface vdsl 0/4
deactivate 1
activate 1 template-index 3
alarm-config 1 1
quit
service-port 3 vlan 50 vdsl mode ptm 0/4/1 multi-service user-vlan untagged inbound
traffic-table index 6 outbound traffic-table index 6
service-port desc 3 description Vlanid:50/vdsl
pitp enable pmode
raio-mode common pitp-pmode
display xdsl vectoring-group 1
display xdsl vectoring-profile 1
xdsl vectoring legacy-cpe activate-policy no-limit
xdsl vectoring enable
save
```

7.8 Vectoring Maintenance and Diagnosis

7.8.1 Common Vectoring Line Faults and Troubleshooting Methods

This topic describes how to rectify a fault using the command line interface (CLI) on a vectoring FTTB/C network.

Context

The fault scenario of the vectoring feature is similar to that of the very-high-speed digital subscriber line 2 (VDSL2) feature. The main difference is that you need to check whether the vectoring function is enabled before rectifying a vectoring fault. Vectoring services affect each other. When the service on a board is unavailable, the services on other boards are affected.

- If the vectoring function is disabled and a service fault occurs, rectify the fault by referring to section "Troubleshooting the FTTB and FTTC Service" in *FTTx Solution Troubleshooting*.
- If the vectoring function is enabled and a fault (**Failure to Access the Internet**, **Low Internet Access Rate**, **Long Time in Switching Programs**, or **Abnormal Interruption of a Multicast Program**) occurs, identify the fault as follows:

a. Check whether the vectoring service board becomes faulty.
b. Check whether the line connecting to the vectoring service board is removed. If the line is removed, the port on the board is deactivated.
c. Check whether the user terminal on the customer premises equipment (CPE) is powered off.
d. Check whether the parameters for the VDSL2 line are set to correct values.

NOTICE

If the VP board becomes faulty, the vectoring function cannot be used (the port cannot be activated using G.993.5). The system automatically activates the port using G.993.2 and then the port can carry common VDSL2 services. If the vectoring function is still required, replace the VP board.

Procedure

Run the **display board** *frameid* command to check whether the vectoring service board becomes faulty.

- If the vectoring service board becomes faulty, replace the board and wait until the fault is rectified.
- If the vectoring service board does not become faulty, go to Step 2.

Step 1 Run the **display board** *frameid/slotid* command to check whether a port on the vectoring service board is deactivated.

- If a port on the vectoring service board is deactivated, reconnect the board using the line and wait until the fault is rectified.
- If no port on the vectoring service board is deactivated, go to Step 3.

Step 2 Check whether a user terminal on the CPE is powered off.

- If a user terminal on the CPE is powered off, power on the terminal and wait until the fault is rectified.
- If no user terminal on the CPE is powered off, go to Step 4.

NOTE

If the preceding three conditions cannot be used to identify a fault or are caused by normal operations (for example, the line is removed to identify another fault on the board, or the user terminal is powered off in a certain period), bit errors may occur on the line in the vectoring group or a user may go offline. At this moment, perform the following steps to check VDSL2 parameter settings of other lines:

Step 3 Run the **display vdsl line-profile** *profile-index* command to check whether the downstream/upstream target signal-to-noise ratio (SNR) margin of the VDSL2 line is set to a correct value.

The default value is 60 (6 dB). Adjust the value based on site requirements. If the noise is large, set the parameter to a large value (80 or 90, or 8 dB or 9 dB).

- If the downstream/upstream target SNR margin of the VDSL2 line is set to a correct value, go to Step 6.
- If the downstream/upstream target SNR margin of the VDSL2 line is set to an incorrect value, go to Step 5.

Step 4 Run the **vdsl line-profile modify** *profile-index* command to set the downstream/upstream target SNR margin of the VDSL2 line. Check whether the fault is rectified.

- If the fault is rectified, go to Step 11.
- If the fault persists, go to Step 6.

Step 5 Run the **display vdsl channel-profile** *profile-index* command to check whether the downstream/upstream minimum impulse noise protection (INP) of the VDSL2 line is set to a correct value.

The default value is 1 (no protection). Adjust the value based on site requirements but do not set it to a value greater than 4 (two symbols).

- If the downstream/upstream minimum INP of the VDSL2 line is set to a correct value, go to Step 8.
- If the downstream/upstream minimum INP of the VDSL2 line is set to an incorrect value, go to Step 7.

Step 6 Run the **vdsl channel-profile modify** *profile-index* command to set the downstream/upstream minimum impulse noise protection of the VDSL2 line. Check whether the fault is rectified.

- If the fault is rectified, go to Step 11.
- If the fault persists, go to Step 8.

Step 7 Run the **display vdsl line-profile** *profile-index* command to check whether the downstream/upstream retransmission function of the VDSL2 line is enabled.

- If the downstream/upstream retransmission function of the VDSL2 line is enabled, go to Step 10.
- If the downstream/upstream retransmission function of the VDSL2 line is disabled, go to Step 9.

Step 8 Run the **vdsl line-profile modify** *profile-index* command to enable the downstream/upstream retransmission function of the VDSL2 line. Check whether the fault is rectified.

NOTE
The retransmission function conflicts with the INP. If both are enabled, only the retransmission function takes effect.

- If the fault is rectified, go to Step 11.
- If the fault persists, go to Step 10.

Step 9 Connect Huawei for technical support.

Step 10 The fault is rectified.

----End

7.8.2 Locating and Troubleshooting of a Vectoring Activation Failure

After vectoring is enabled globally, ports in a vectoring group are activated in common mode. When vectoring is invalid on the port, see this topic to locate and rectify the fault.

Context

After vectoring is enabled globally, run the **display line operation** command to query the port activation mode of the vectoring group. If **Standard in port training** in command output is **G.993.5**, vectoring takes effect on ports. If it is not **G.993.5**, vectoring does not take effect on ports.

The possible causes are:

- Vectoring is disabled.
- The CPE does not support G.993.5.
- Upstream and downstream crosstalk cancellations are not enabled.
- Bandplan division encounters a compatibility problem.

Procedure

Run the **display xdsl vectoring config** command to check whether vectoring is enabled globally. Ports can be activated in vectoring mode only after vectoring is enabled globally. If vectoring is not enabled globally, run the **xdsl vectoring** command to enable it globally.

Step 1 Run the **display inventory cpe** command to check whether the transmission mode capability set of the CPE supports G.993.5. Ports can be activated in vectoring mode only when the transmission mode capability set of the CPE supports G.993.5.

```
huawei(config-if-vdsl-0/4)#display  inventory cpe 2
  ------------------------------------------------------------------------
  G.994.1 vendor ID             : 0xB5004244434D0000
    G.994.1 country code         : 0xB500
    G.994.1 provider code        : BDCM
    G.994.1 vendor info          : 0x0000
  System vendor ID              : 0xB5004244434D0000
    System country code          : 0xB500
    System provider code         : BDCM
    System vendor info           : 0x0000
  Version number                : A2pv6C037g
  Version number(octet string)    : 0x413270763643303337670000000000000
  Vendor serial number           : -
  Self-test result              : PASS
  Transmission mode capability    :
    G.992.1(Annex A)             G.992.3(Annex A)
    G.992.5(Annex A)             G.993.2(Annex A/B/C)
    G.993.5                     Full G.993.5 friendly            //Supports G.993.5
mode//
 ------------------------------------------------------------------------
```

Step 2 Run the **display xdsl vectoring-profile** command to check whether upstream and downstream crosstalk cancellation are enabled. If they are disabled, run the **xdsl vectoring-profile modify** command to enable them.

Step 3 Run the **display xdsl vectoring line-info** command to check whether bandplan configured on the profile bound to the port is compatible with that configured globally. If **BandPlan Compatible** in the command output is **Y**, the bandplans are compatible. If it is **N**, the bandplans are not compatible. Run the **xdsl vectoring bandplan-type** command to select the bandplan configured globally that is compatible with that on the profile bound to the port, and activate the port again.

Step 4 Connect Huawei for technical support.

----End

7.8.3 N2510 Vectoring O&M

Huawei N2510 vectoring solution provides vectoring line evaluation, fault diagnosis, and line optimization, helping carriers to visualize vectoring benefits, reduce O&M costs, and improve line stability, thereby improving end users' satisfaction degree.

For details, see *At a Glance of N2510 Features - Vectoring O&M*.

7.9 Vectoring Reference Standards and Protocols

Standard/Protocol
ITU-T G.993.5: Self-FEXT cancellation (vectoring) for use with VDSL2 transceivers
WT-249: Testing of Self-FEXT Cancellation (vectoring)
ITU-T G.993.2: Very high speed digital subscriber line transceivers 2 (VDSL2)
ITU-T G.994.1: Draft Amendment 8 to Recommendation ITU-T G.994.1
ITU-T G.994.1: Mandatory tone set for HPE17 and HPE30, and codepoints in support of Recommendations ITU-T G.993.5 and ITU T G.998.4
ITU-T G.997.1: Management of ITU-T G.998.4, G.993.5 and receiver referred virtual noise of ITU-T G.993.2
ITU-T G.997.1: Physical layer management for digital subscriber line (DSL) transceivers - 2 Amendment 4

7.10 Vectoring Acronyms and Abbreviations

Acronyms and Abbreviations	Full Name
ADSL2+	Asymmetric Digital Subscriber Line 2 Plus
AFE	Analogue Front End
AN	Access Node
CO	Central Office
CP	Customer Premises
CPE	Customer Premises Equipment
DLM	Dynamic Line Management

Acronyms and Abbreviations	Full Name
DSL	Digital Subscriber Line
DSE	Disorderly Shutdown Event
DSLAM	DSL Access Multiplexer
DSM	Dynamic Spectrum Management
EMS	Element Management System
ITU	International Telecommunication Union
OPEX	Operational Expenditure
ERB	Error Report Block
ES	Error Samples
FTTB	Fiber to the Building
FTTC	Fiber to the Curb
PSD	Power Spectral Density
ME	Management Entity
MIMO	Multiple Input Multiple Output
NDR	Net Data Rate
FDM	Frequency-division multiplexing
PMD	Physical Medium Dependent
L2+	Ethernet Layer 2 and Above
RT	Remote Terminal
ETR	Expected Throughput
VCE	Vectoring Control Entity
VDSL2	Very High Speed Digital Subscriber Lines 2
VME	Vectoring Management Entity
VCU	Vectoring Control Unit
VP	Vectoring Processor
VN	Virtual Noise
VTU	VDSL Transceiver Unit

8 SHDSL Access

About This Chapter

SHDSL is an xDSL access technology, just like ADSL and VDSL. SHDSL provides the symmetric upstream and downstream rates.

8.1 ATM SHDSL Access

This topic describes the definition, purpose, specifications, and limitations of ATM SHDSL access feature. It also provides the glossary and the acronyms and abbreviations related to the ATM SHDSL access feature.

8.1.1 Introduction

Definition

SHDSL is an xDSL access technology, just like ADSL and VDSL. SHDSL provides the symmetric upstream and downstream rates.

The symmetric upstream and downstream rates of ATM SHDSL determine that bi-directional rates of the supported service must be basically the same. In addition, ATM SHDSL features a longer transmission distance. Hence, ATM SHDSL can be widely used.

Purpose

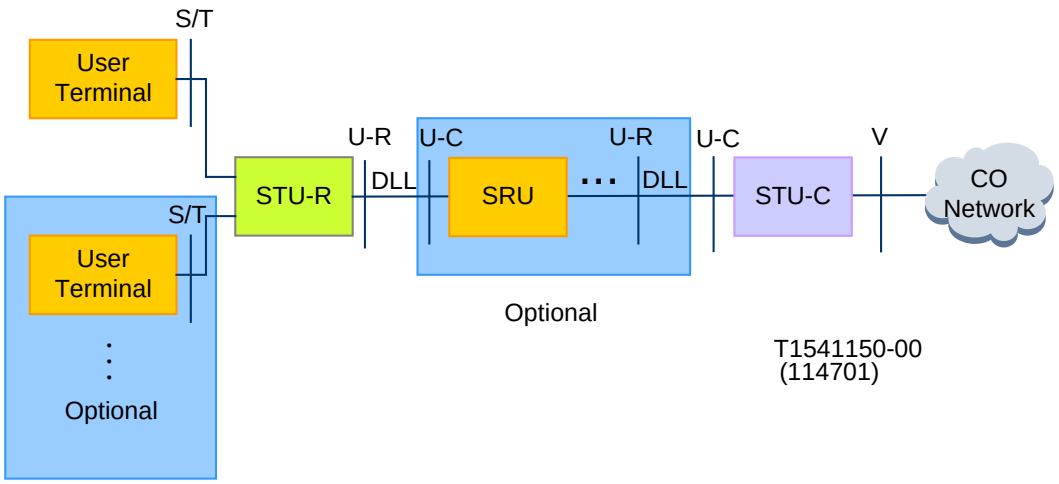
ATM SHDSL provides symmetric broadband access services for subscribers to meet the requirement for high downstream rate from SOHO subscribers. ATM SHDSL applications are similar to ADSL applications and the ATM SHDSL and ADSL applications are mutually complementary.

8.1.2 Principle

Typical Application Model

The SHDSL operating principle is based on the G.991.2 (2001) standard.

Figure 8-1 Typical application model of SHDSL



One SHDSL system consists of an STU-C, an STU-R, and a subscriber terminal. Multiple repeaters can be added to the line between the STU-C and the STU-R.

- The STU-C provides service ports at the central office.
- The STU-R provides subscriber ports for connecting to multiple subscriber terminals.
- The SHDSL repeater unit (SRU) refers to the repeater. In ultra-long distance transmission, it recovers signals and re-transmits signals to increase the transmission distance.

The MA5600T/MA5603T/MA5608T does not support ATM SHDSL repeaters.

Terminal Model

The SHDSL terminal model consists of the following parts:

- PDM module
  - The PDM module implements functions such as: Regular code element generation and recovery, coding/decoding, modulation/demodulation, echo control, linear equalization, and link start
  - SHDSL mainly uses the trellis coded pulse amplitude modulation (TC-PAM) technology.
- PMS-TC module

  The PMS-TC module implements functions such as: framing, frame synchronization scrambling, and descrambling
- TPS-TC module

  The TPS-TC module implements functions such as: mapping and encapsulation of data frames, multiplexing and demultiplexing, timing alignment of multiple subscriber data channels
- I/F interface of the device at the central office
  - It mainly provides the ATM port.
  - The ATM port is used for transmitting ATM cells over the ATM network, or according to the carried packets, transmitting Ethernet packets encapsulated by the SAR module or E1/V3.5 signals over the Ethernet network.

- I/F interface of the device on the subscriber side

  It corresponds to the I/F interface of the device at the central office. In general, the I/F interface is used for providing Ethernet ports or E1/V.35 ports.

When the MA5600T/MA5603T/MA5608T uses the SHLB board, the TC-PAM encoding technology is shown as the following table.

Table 8-1 TC-PAM encoding technology

Compliant Standards	Describes...
SHDSL	$R = n \times 64 + (i) \times 8$, $3 \leq n \leq 36$ and $0 \leq i \leq 7$ (192 kbit/s to 2312 kbit/s)

The SHLB board of the MA5600T/MA5603T/MA5608T is based on ATM. The board provides the Ethernet port (for broadband access) or E1/V.35 port (for private line access) for connecting subscriber terminals. In the upstream direction, the board is connected to the metropolitan area network (MAN) through the upstream board.

8.1.3 IMA Introduction

IMA Overview

Inverse multiplexing over ATM (IMA) allows a sender to break up the ATM cell flows and distributes the cells over multiple low-speed links, and allows a receiver to recombine the cells into the cell flows. IMA enables the transmission of ATM cells over existing links (especially 2 Mbit/s links).

The IMA technology includes multiplexing and demultiplexing of ATM cells. The functional group that performs the multiplexing and demultiplexing is called an IMA group. An IMA group terminates at the end of each IMA virtual connection.

In practice, users can use the IMA technology to transmit services over one or multiple G.SHDSL links based on desired bandwidths.

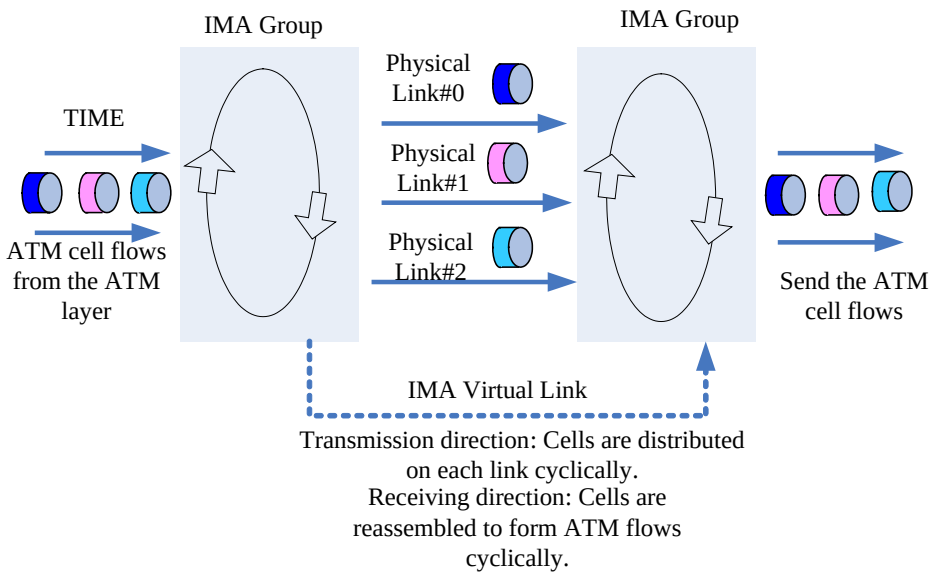
IMA Principles

ATM cells are transmitted over links using the Round Robin distribution mechanism. This mechanism allows each separate cell to be cyclically sent over links. An IMA group periodically sends IMA Control Protocol (ICP) cells to define IMA frames. ICP cells enable the receiver to reconstruct ATM cell flows. Based on the arrival time of IMA frames, the receiver can detect and adjust the link differential delay to remove the cell delay variation (CDV) imported by ICP cells.

The sender sends cells consecutively. If no ATM layer cell can be sent in an IMA frame between ICP cells, the IMA sender adds filler cells to ensure consecutive cells. These filler cells will be discarded by the IMA receiver.

Figure 8-2 shows the transmission of ATM cells in an IMA group.

Figure 8-2 Transmission of ATM cells in an IMA group



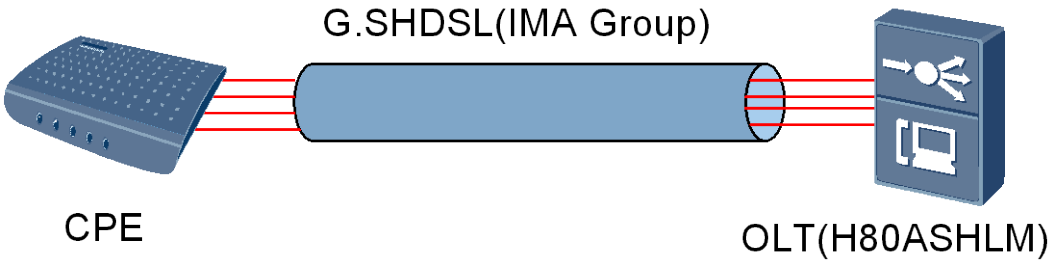
IMA System

As shown in Figure 8-3 one or multiple physical links are connected between the customer premises equipment (CPE) and the OLT (MA5600T).

- A single G.SHDSL link can be used to transmit services.
- Multiple G.SHDSL links can be bonded to form an IMA group to transmit services if a single G.SHDSL link cannot provide the desired bandwidth.

An IMA group is a logical link multiplexing one or multiple low-speed links. It provides a high bandwidth and supports high-speed ATM cell flows. The bandwidth of an IMA group is approximately the sum of the bandwidths of all member links. IMA technology is flexible to use and cost-effective.

Figure 8-3 IMA system



Troubleshooting Procedure

- Run the **display port state** command to check whether an SHDSL port is activated (check whether the port status is Activated).

- Run the **display ima group run-status** command to check whether the near end and far end of the IMA group are functional (check whether the status of **NE group state** and **FE group state** is start-up).
- Run the **display ima link run-status** command to check whether the IMA link is functional at both near end and remote end (check whether the status of **NE Tx link state**, **NE Rx link state**, **FE Tx link state**, and **FE Rx link state** is Active).

The following events reflect IMA Troubleshooting. If any event is reported, rectify the fault.

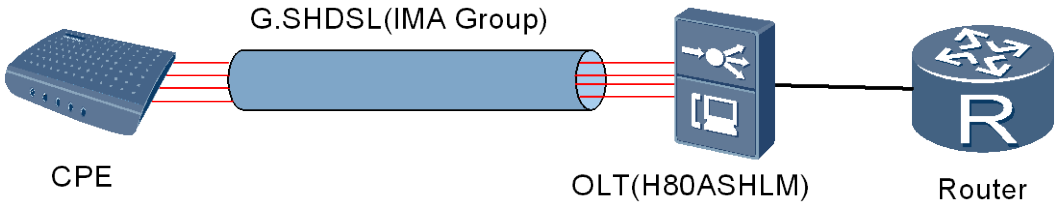
Event	Name
0x11300100	The status of the IMA group changes
0x11300006	The bandwidth of the IMA group changes
0x11300101	The status of the IMA link changes

8.1.4 Configuration Examples of IMA

Networking

Figure 8-4 shows the MA5600T/MA5603T/MA5608T networking.

Figure 8-4 IMA service networking

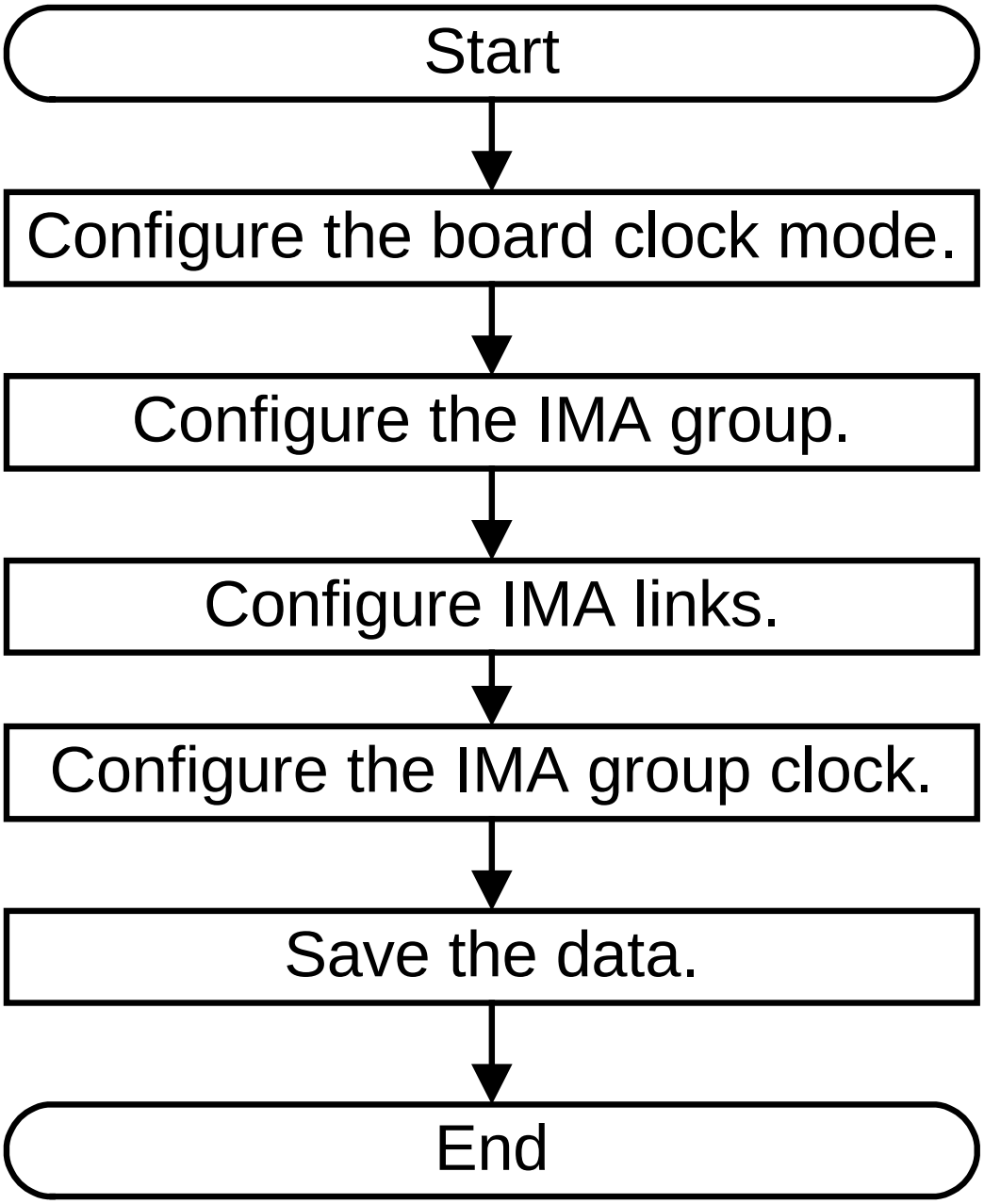


Configuration Flowchart

The MA5600T/MA5603T/MA5608T provides IMA interfaces using an H80ASHLM board for connecting to a remote device.

Figure 8-5 shows the flowchart for configuring the IMA service.

Figure 8-5 Flowchart for configuring the IMA service



Data Plan

Table 8-2 Data Plan

Parameter	MA5600T/MA5603T/MA5608T	CPE
IMA group ID	0	0
IMA version	version1.1	version1.1

Parameter	MA5600T/MA5603T/MA5608T	CPE
IMA ID	0	0
TX clock mode	CTC	ITC
IMA link ID	0-3	0-3
Clock mode	system	line

NOTE
In IMA service networking, CPE configurations vary depending on CPE types. This document describes only the configurations on the MA5600T/MA5603T/MA5608T.

Procedure

Run the **display board** command to verify that the H80ASHLM board is functional

```
huawei(config)#display board 0
 ------------------------------------------------------------------------------
 SlotID  BoardName  Status         SubType0 SubType1   Online/Offline
 ------------------------------------------------------------------------------
 0
 1
 2      H806GPBD   Failed                            Online
 3      H802GPBD   Normal
 4      H802SHLB   Normal
 5      H805GPFD   Normal
 6      H803GPFD   Normal
 7      H807GPBD   Normal
 8      H802EPBD   Failed                            Offline
 9
 10     H801SCUN   Active_normal
 11
 12     H802SHGM   Normal
 13     H80ASHLM   Normal    //H80ASHLM board is normal
 14
 15     H802EPBD   Normal
 16     H805GPBD   Normal
 17     H805ADPD   Normal
 18     H801OPFA   Normal
 19     H801GICG   Normal
 20
 21
 22
 ------------------------------------------------------------------------------
```

Step 1 Run the **interface shl** command to enter SHDSL mode, and run the **set clockmode** command to configure the board clock to lock the system clock. The system clock of the MA5600T/MA5603T/MA5608T features high precision. Therefore, the board clock locks the system clock of the MA5600T/MA5603T/MA5608T in practice.

```
huawei(config)#interface shl 0/13
huawei(config-if-shl-0/13)#set clockmode
{ status<E><free-run,system> }:system

  Command:
        set clockmode system
  The new clock mode will not take effect until the port is activated again.
  Are you sure to set clock mode? (y/n)[n]:y
```

NOTICE

- If the port is deactivated, configure the network clock mode. Then, run the activate command to activate this port.
- If the port is activated, run the deactivate command to deactivate it. Then, configure the network clock mode and run the activate command to activate this port.

Step 2 Run the **ima group add** command to add IMA group 17 containing four links (0-3). Then, set the CTC mode for the clock.

```
huawei(config-if-shl-0/13)#ima group add
{ groupIndex<U><16,31> }:17
{ version<E><version1.0,version1.1> }:version1.1
{ minTxLinks<U><1,16> }:1
{ minRxLinks<U><1,16> }:1
{ clock<E><itc,ctc> }:ctc      //Set the CTC mode on the MA5600T
{ imaid<U><0,255> }:0         //Ensure that the IMA ID is the same on the MA5600T and
the CPE
{ framelength<E><32,64,128,256> }:128
{ alpha_value<U><1,2> }:2
{ beta_value<U><1,5> }:2
{ gamma_value<U><1,5> }:1

  Command:
        ima group add 17 version1.1 1 1 ctc 0 128 2 2 1
```

Step 3 Run the **ima link add** command to Add the four links to IMA group 0. If the line is functional, the bandwidth of the IMA group is increased after the links are added successfully. To query link status, run the **display ima link run-status** command.

```
huawei(config-if-shl-0/13)#ima link add
{ groupIndex<U><16,31> }:17
{ linkid<U><0,15> }:0

  Command:
        ima link add 17 0

huawei(config-if-shl-0/13)#ima link add
{ groupIndex<U><16,31> }:17
{ linkid<U><0,15> }:1

  Command:
        ima link add 17 1
```

```
huawei(config-if-shl-0/13)#ima link add
{ groupIndex<U><16,31> }:17
{ linkid<U><0,15> }:2

  Command:
         ima link add 17 2

huawei(config-if-shl-0/13)#ima link add
{ groupIndex<U><16,31> }:17
{ linkid<U><0,15> }:3

  Command:
         ima link add 17 3
```

Step 4 Run the **ima group mode clockmode** command to configure the clock of IMA group 17 on the MA5600T/MA5603T/MA5608T to lock the system clock.

```
huawei(config-if-shl-0/13)#ima group mode
{ clockmode<K>|crc4-multiframe<K>|scramble<K> }:clockmode
{ all<K>|groupIndex<U><16,31> }:17
{ line<K>|system<K> }:system    //Lock the system clock of the MA5600T

  Command:
         ima group mode clockmode 17 system
```

NOTICE

The CPE clock must be synchronized with the MA5600T/MA5603T/MA5608T clock. The MA5600T/MA5603T/MA5608T clock is the master clock, and the CPE clock is the slave clock.

Step 5 Save the data.

```
huawei(config-if-shl-0/13)#quit
huawei(config)#save
```

----End

8.1.5 Reference

The following lists the reference documents of this feature:

- ITU-T Recommendation G.991.2 Annex A and Annex F.
- ITU-T Recommendation G.991.2 Annex B and Annex G.
- RFC 4319: Definitions of Managed Objects for High Bit-Rate DSL - 2nd generation (HDSL2) and Single-Pair High-Speed Digital Subscriber Line (SHDSL) Lines

8.2 EFM SHDSL Access

This topic describes the definition, purpose, specifications, and limitations of EFM SHDSL access feature. It also provides the glossary and the acronyms and abbreviations related to the EFM SHDSL access feature.

8.2.1 Introduction

Definition

SHDSL is an xDSL access technology, just like ADSL and VDSL. SHDSL provides the symmetric upstream and downstream rates.

EFM SHDSL integrates the advantages of the SHDSL technology and the ADSL technology. That is, EFM SHDSL can provide traditional voice service and high rate Internet access service over common twisted pairs to meet the requirements for high definition TV service and VoD service from subscribers, which suit the last mile access for broadband to the campus.

Purpose

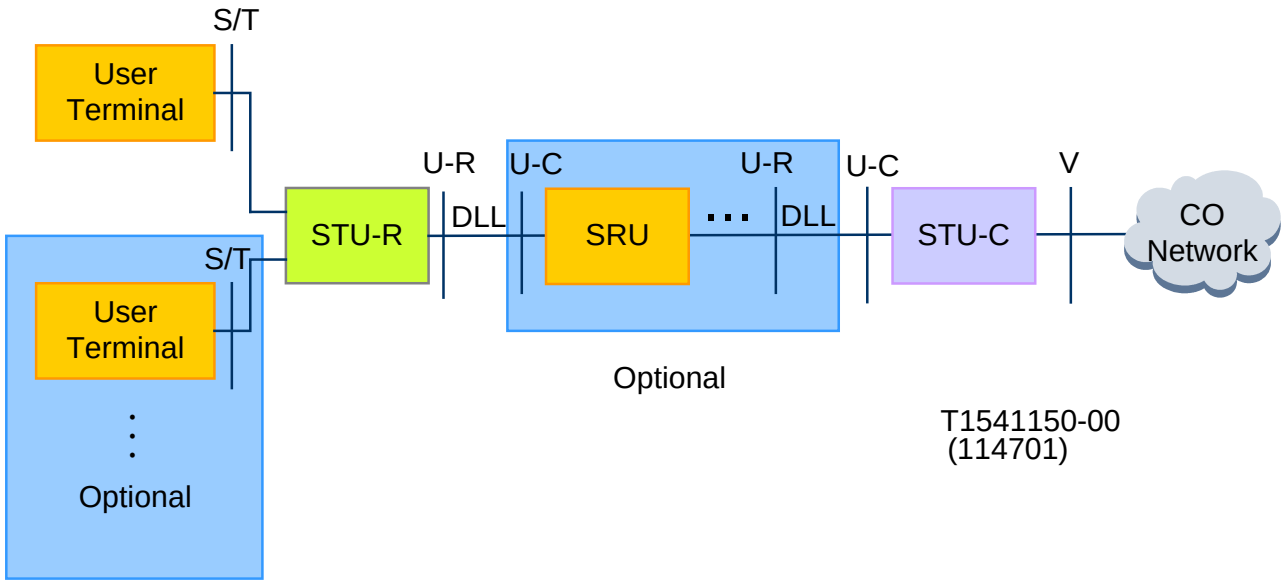
The utilization ratio of the EFM access service is high when the activation rates of the ATM and EFM access services are the same. Hence, if the subscriber terminal supports ATM and EFM SHDSL access services simultaneously, the EFM SHDSL access service is preferred.

8.2.2 Principle

Typical Application Model

The SHDSL operating principle is based on the G.991.2 (2001) standard.

Figure 8-6 Typical application model of SHDSL



One SHDSL system consists of an STU-C, an STU-R, and a subscriber terminal. Multiple repeaters can be added to the line between the STU-C and the STU-R.

- The STU-C provides service ports at the central office.
- The STU-R provides subscriber ports for connecting to multiple subscriber terminals.
- The SHDSL repeater unit (SRU) refers to the repeater. In ultra-long-distance transmission, it recovers signals and re-transmits signals to extend the transmission distance.

Terminal Model

The SHDSL terminal model consists of the following parts:

- PDM module
  - The PDM module implements functions such as: Regular code element generation and recovery, coding/decoding, modulation/demodulation, echo control, linear equalization, and link start
  - SHDSL mainly uses the trellis coded pulse amplitude modulation (TC-PAM) technology.
- PMS-TC module

  The PMS-TC module implements functions such as: framing, frame synchronization scrambling, and descrambling
- TPS-TC module

  The TPS-TC module implements functions such as: mapping and encapsulation of data frames, multiplexing and demultiplexing, timing alignment of multiple subscriber data channels
- I/F interface of the device at the central office
  - Providing ATM ports or circuit interfaces
  - The ATM port is used for transmitting ATM cells over the ATM network, or according to the carried packets, transmitting Ethernet packets encapsulated by the SAR module or E1/V3.5 signals over the Ethernet network or E1 links.
  - The circuit interface is used for transmitting E1 or V.35 signals directly through the time division multiplexing (TDM) network.
- I/F interface of the device on the subscriber side

  It corresponds to the I/F interface of the device at the central office. In general, the I/F interface is used for providing Ethernet ports (for delivering ATM cells processed by the SAR module) or E1/V.35 ports.

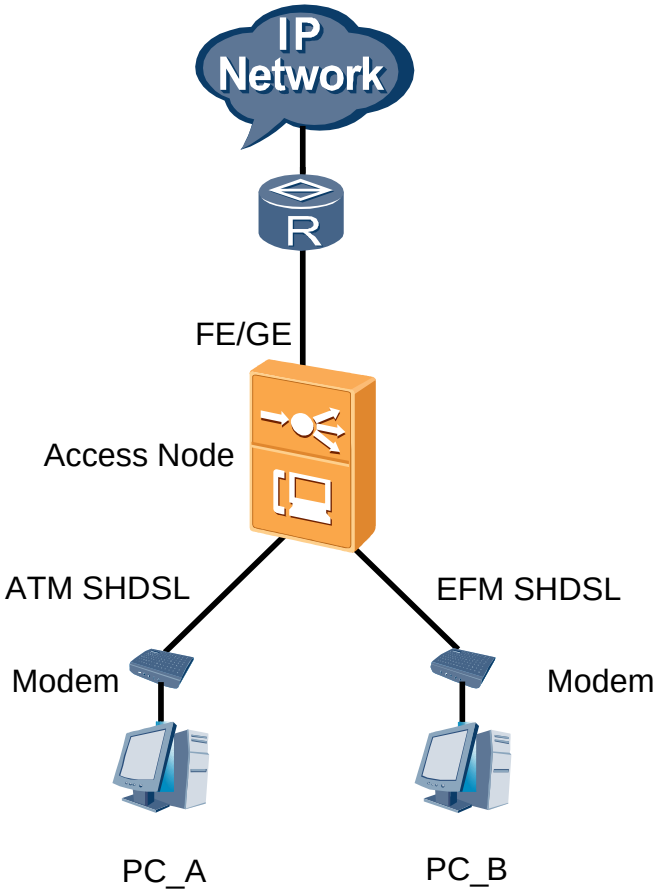
Table 8-3 TC-PAM encoding technology

Compliant Standards	Describes...
SHDSL	$R = n \times 64 + (i) \times 8$, $3 \leq n \leq$ 89 and $0 \leq i \leq 7$ (192 kbit/s to 5696 kbit/s)

Typical Networking Application

The Figure 8-7 shows the typical networking application of EFM SHDSL.

Figure 8-7 Typical networking application of EFM SHDSL



8.2.3 Reference

The following lists the reference documents of this feature:

- ITU-T Recommendation G.991.2 Annex A and Annex F.
- ITU-T Recommendation G.991.2 Annex B and Annex G.
- RFC 4319: Definitions of Managed Objects for High Bit-Rate DSL - 2nd generation (HDSL2) and Single-Pair High-Speed Digital Subscriber Line (SHDSL) Lines

8.3 TDM SHDSL Feature

8.3.1 Introduction

Definition

Single-pair high-speed digital subscriber line (SHDSL), defined by ITU-T (such as ITU-T G.991.2), is a data transmission technology over twisted pairs to transmit voice, data, and video signals.

TDM SHDSL is a mode to transmit TDM signals through SHDSL.

As the transmission mode varies, the device provides different types of upstream ports. Specifically, the TDM-E1-G.703 electrical port is used by the device for the TDM transmission system; the ATM-STM-1 optical port is used by the device for the ATM transmission system. Similarly, the user-side CPE also provides different types of data ports to adapt to different transmission modes. Specifically, for the TDM transmission system, the CPE generally provides the TDM-V.35 or E1-G.703 port; for the ATM transmission system, the CPE generally provides the ATM-FR-V.35, 10/100Base-T Ethernet, or ATM-CE-V.35 (or E1-G.703) port.

Purpose

TDM SHDSL provides the TDM-V.35 or E1-G.703 port. Compared with the V.35 and E1 cables, SHDSL has an advantage of farther transmission distance; therefore, SHDSL can extend the reach of DDN nodes over abundant twisted pair resources.

TDM SHDSL achieves E1 transmission and access over subscriber cables at "last two miles" and at the same time carries various services of N x 64 kbit/s. Hence, TDM SHDSL makes possible the broadband private line access for users over the existing transmission network resources.

Benefit

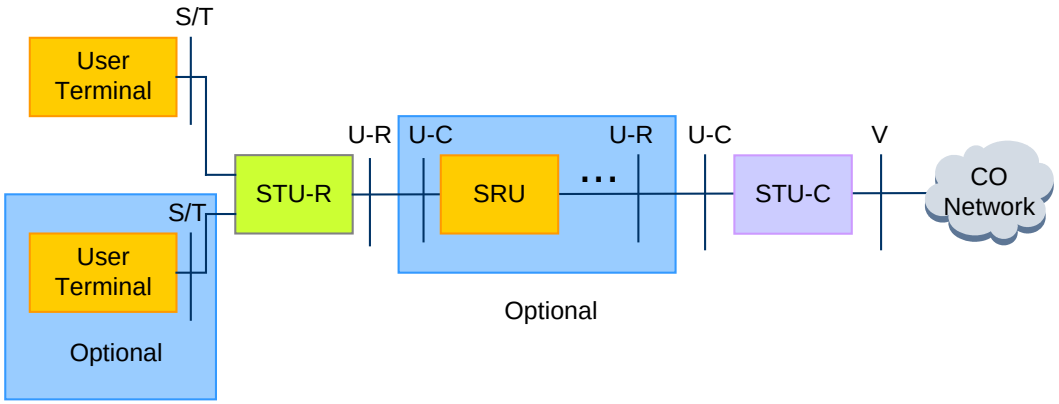
The abundant twisted pair resources can be utilized to achieve the long-distance access of the circuit emulation equipment with the E1 or V.35 port, thereby reducing the consumption of copper wire resources.

8.3.2 Principle

Basic Principle

Based on the G.991.2 (2001) standard, the SHDSL system consists of an SHDSL transceiver unit at the Central Office (STU-C), an SHDSL transceiver unit at the Remote End (STU-R), and a user terminal. Between STU-C and STU-R, there may be several SHDSL regenerator units (SRUs), as shown in Figure 8-8.

Figure 8-8 Typical application model of SHDSL



SRU: SHDSL Regenerator Unit STU: SHDSL Transceiver Unit

STU-C: STU at the Central Office STU-R: STU at the Remote End

- The STU-C provides the service ports on the CO side.
- The STU-R provides the user ports. One STU-R can be connected to multiple user terminals.
- SRUs are used in ultra-distance transmission and it recovers signals and re-transmits signals to increase the transmission distance.

STU-Cs are generally placed in a centralized manner and provide network-side upstream ports to form the DSLAM equipment. According to the varying transmission mode in the system, the DSLAM provides different upstream ports.

- In the case of the TDM transmission system, the DSLAM generally provides the TDM-E1-G.703 electrical port.
- In the case of the ATM transmission system, the DSLAM generally provides the ATM-STM-1 optical port.

The STU-R and user-side data port form the user-side CPE. Similarly, the CPE provides different user-side ports to meet the requirements of the varying transmission modes.

- In the case of the TDM transmission system, the CPE generally provides the TDM-V.35 or E1-G.703 port.
- In the case of the ADM transmission system, the CPE generally provides the ATM-FR-V.35, 10/100Base-T Ethernet port, or ATM-CE-V.35 (or E1-G.703) port.

NOTE

In the case of the TDM transmission system, the MA5600T/MA5603T/MA5608T supports only the TDM-E1-G.703 electrical port for upstream transmission and only TDM SHDSL (E1) on the user side.

In the case of the ATM transmission system, because the IP network is a mainstream network, the MA5600T/MA5603T/MA5608T does not support the ATM-STM-1 optical port for upstream transmission but the MA5600T/MA5603T/MA5608T supports ATM access.

Working Mode

The H802EDTB board can work in the VOICE mode and SAToP mode.

In the case of TDM SHDSL in the VOICE mode, the H802EDTB board needs to be configured with the working sub-mode: service mode, transparent transmission mode or PRA-v3 mode.

- Service mode

  Each G.SHDSL port and E1 port are independent ports, on which the SPC, VoIP ISDN PRA service (IP upstream), port rate, or port mode can be configured.

- Transparent transmission mode

  The H802EDTB board automatically connects the Nth SHDSL line with the Nth E1 line to transparently transmit the 2M data. The E1 port is in the UNFRAME format. The clock locks the Nth E1 line clock. Therefore, every E1 line has its independent clock. In the transparent transmission mode, the SPC and PRA services cannot be configured.

- PRA-v3 mode

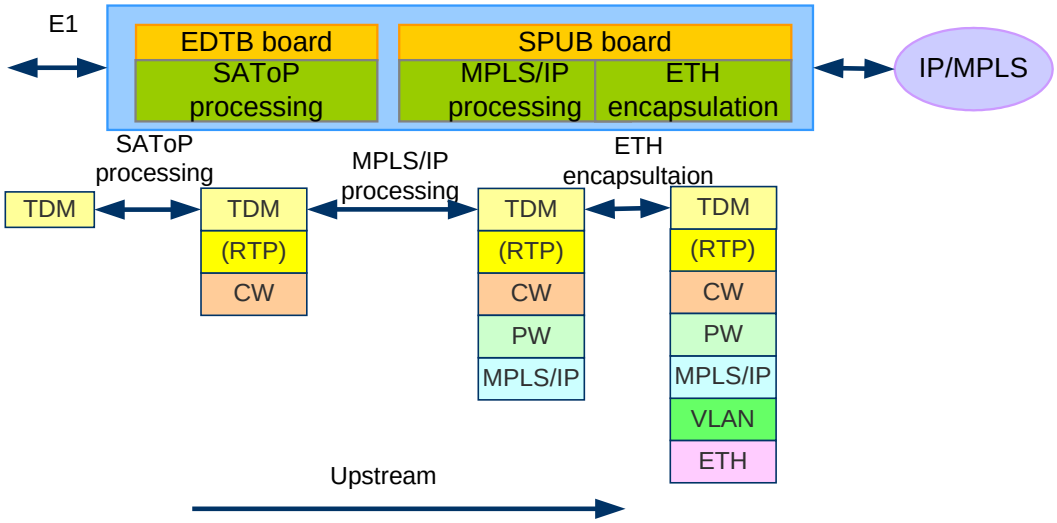
  In this mode, the E1 ports and the G.SHDSL ports on the H802EDTB board are one-to-one mapping (ports 0-15 and ports 16-31 are one-to-one mapping) to implement the ISDN PRA service (E1 upstream), and to receive and process the **loopback1** command sent from the V3 reference point.

- Data mode

Indicates the TDM data mode. It is used for the G.704 data service scenario. In this mode, the E1 ports and G.SHDSL ports on the H802EDTB board have a one-to-one mapping relation (E1 ports 0-15 and G.SHDSL ports 16-31 have a one-to-one mapping relation) to implement data transmission. The cyclic redundancy check (CRC) function is enabled on both E1 and G.SHDSL ports to ensure data transmission reliability.

In the case of TDM SHDSL in the SAToP mode, the MA5600T/MA5603T/MA5608T supports E1 access, and also supports SAToP encapsulation and processing of E1 service. Figure 8-9 shows the service processing flow.

Figure 8-9 Processing flow of TDM PWE3 service in E1 access



- Packing/Unpacking of SAToP packets

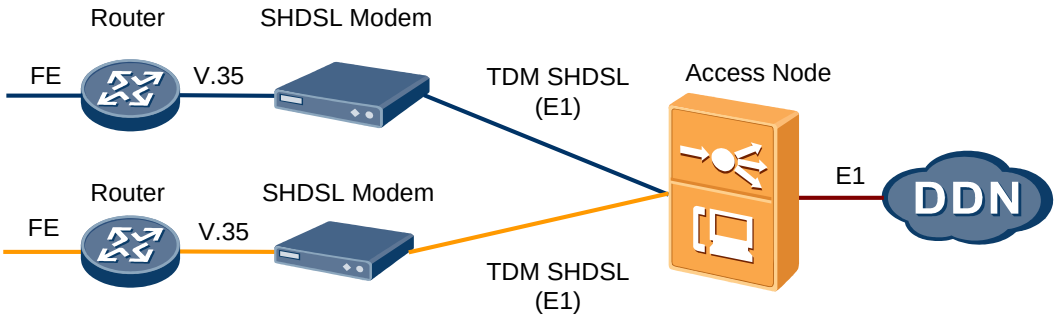
  The MA5600T/MA5603T/MA5608T packs E1 data in the SAToP format, and adds the control word and RTP header (optional in the MPLS mode) to the SAToP packets.
- Encapsulation of MPLS labels
  - The MA5600T/MA5603T/MA5608T adds/deletes the MPLS labels, and maps inner labels to user circuits.
  - In the MPLS+MPLS encapsulation, the outer LSP label is used for transmitting the packet over an MPLS network; in the IP+MPLS encapsulation, the outer IP address is used for transmitting the packet over an IP network. The inner label is used for mapping to a user circuit.
  - The inner PW tunnel is a bidirectional MPLS tunnel that carries TDM data. A PW label can be statically configured or dynamically created through protocol (LDP).
  - The outer tunnel can be MPLS-encapsulated or IP-encapsulated. In the case of MPLS encapsulation, the outer MPLS tunnel can be statically configured or dynamically created through protocol (LDP or RSVP-TE). In the case of IP encapsulation, the outer IP tunnel can be statically configured.
- Ethernet processing: In the upstream direction, the ETH header is encapsulated to the packet label header, and then the packet is transmitted through the upstream port on the control board.
  - The upstream VLAN of the TDM PWE3 packet is a service VLAN, which is the VLAN of the corresponding upstream port.

– The Layer 3 interface MAC address is filled in as the source MAC address of the TDM PWE3 upstream packet, and the MAC address of the next-hop interface (this MAC address can be learned through ARP) is used as the destination MAC address.

8.3.3 Narrowband Data Private Line Service Applications

The narrowband data private line service is mainly demonstrated in expanding the reach of DDN nodes. TDM SHDSL for expanding the reach of DDN nodes is a mainstream method supported by the integrated access equipment to provide the DDN service. On the CO side, the integrated access equipment connects to the DDN node through E1; on the user side, the TDM-capable SHDSL modem provides the TDM SHDSL (E1) port to implement N x 64 kbit/s private line access and at the same time achieves private line interconnection by supporting the V.35-capable router, as shown in Figure 8-10.

Figure 8-10 Narrowband data private line service applications

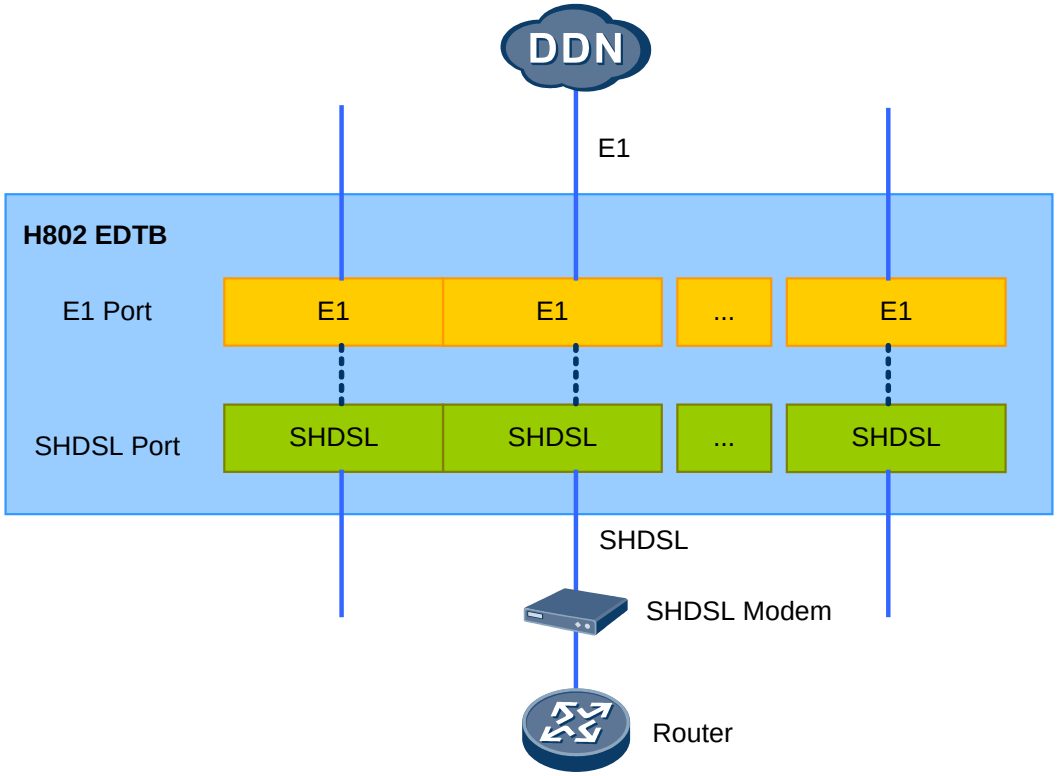


The MA5600T/MA5603T/MA5608T connects to the DDN node in the following two ways:

- Transparent transmission
- Aggregation

Figure 8-11 shows how the MA5600T/MA5603T/MA5608T connects to the DDN node in the transparent transmission mode: The H802EDTB board connects upstream to the DDN network through E1 and connects downstream to the SHDSL modem through SHDSL.

Figure 8-11 Connection to the DDN (in the transparent transmission mode)



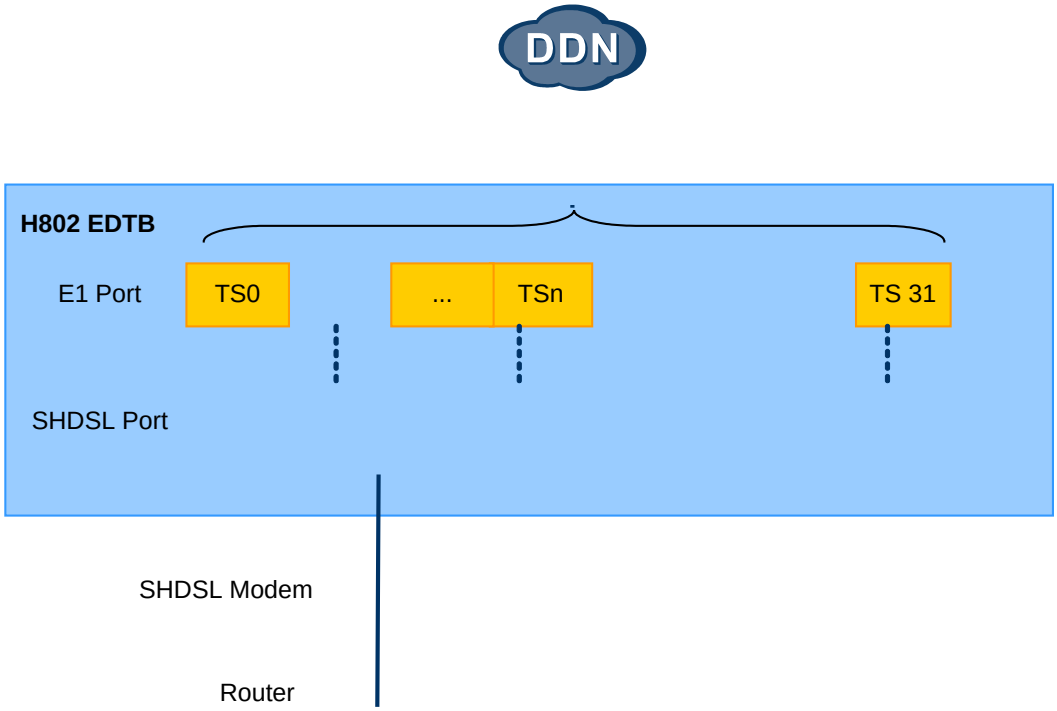
The working sub-mode of the H802EDTB board of the MA5600T/MA5603T/MA5608T is set to the transparent transmission mode. In this mode, the H802EDTB board automatically maps E1 ports 0-15 to SHDSL ports 16-31 to transparently transmit data.

NOTE

In addition, the clock source for every E1 port on the H802EDTB board comes from the E1 line clock and the clock source for an SDHSL port keeps synchronized with its corresponding E1 port.

Figure 8-12 shows how the MA5600T/MA5603T/MA5608T connects to the DDN node in the aggregation mode: The H802EDTB connects upstream to the DDN network through E1 and connects downstream to the SHDSL modem through SHDSL.

Figure 8-12 Connection to the DDN (in the aggregation mode)



An SHDSL port supports only framed N x 64 kbit/s, that is, the SHDSL modem still sends 32 x 64 kbit/s to the equipment (certain timeslots of the 32 timeslots may not carry data because N may be smaller than 32). In this way, The H802EDTB board aggregates certain timeslots in 32 x 64 kbit/s for multiple SHDSL ports and then sends them upstream to the DDN.

NOTE

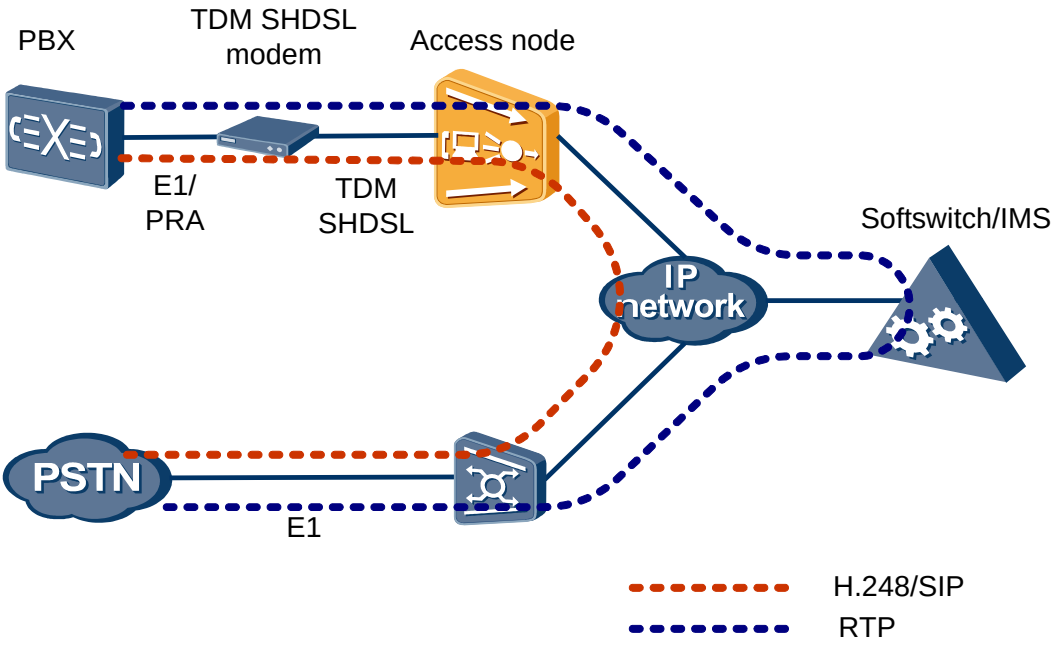
That is, N x 64 kbit/s is input to the SHDSL modem and the modem outputs E1 frames with 32 timeslots. The equipment aggregates certain timeslots of multiple E1 frames into a same E1 port and then sends them upstream to the DDN.

The working sub-mode of the H802EDTB board of the MA5600T/MA5603T/MA5608T is set to the service mode. In addition, the frame format of the E1 and SHDSL ports are configured to UNFRAME, and SPCs are set up for timeslots between N x 64 kbit/s for multiple SHDSL ports and E1 ports. This achieves the aggregation of multiple N x 64 kbit/s into E1, that is, timeslot channels of different lines are multiplexed to the same E1 upstream port, thereby saving E1 resources.

8.3.4 PRA Carrying Applications

Figure 8-13 shows the long-distance access of the PBX to the IP network for carrying the PRA service.

Figure 8-13 PRA carrying applications



- The PBX provides E1 in the upstream direction.
- The SHDSL modem implements the E1-to-SHDSL conversion and connects upstream through SHDSL to the SHDSL port on the H802EDTB board of the MA5600T/MA5603T/MA5608T.
- The MA5600T/MA5603T/MA5608T connects upstream to the IP network.
- The working mode of the H802EDTB board of the MA5600T/MA5603T/MA5608T is configured to the service mode.
- The signaling mode of the SHDSL port is configured to CCS. In addition, the D channel signaling of the PRA is transmitted in timeslot 16 and timeslot 0 is used for frame synchronization.

By using SHDSL, the MA5600T/MA5603T/MA5608T provides long-distance transmission to implement long-distance access of the MA5600T/MA5603T/MA5608T and PBX.

8.3.5 Reference Standards and Protocols

The reference standards and protocols of the TDM SHDSL feature are as follows:

- G.991.2 Annex A and Annex F: Standards applicable for North America
- ITU-T G.991.2 Annex B and Annex G: Standards applicable for European
- RFC4319 Definitions of Managed Objects for High Bit-Rate DSL - 2nd generation (HDSL2) and Single-Pair High-Speed Digital Subscriber Line (SHDSL) Lines

8.4 Configuration SHDSL

SHDSL service configuration includes SHDSL profile configuration and SHDSL user port configuration. This topic describes the detailed configuration methods and procedures.

8.4.1 Configuring SHDSL Profiles

This topic describes how to configure the SHDSL line profile and alarm profile.

Context

The SHDSL line profile and alarm profile can be directly bound to an SHDSL port.

Table 8-4 lists the default SHDSL profiles.

Table 8-4 Default SHDSL profiles

Parameter	Default Setting
SHDSL line profile	Profile IDs: 1, 100, 101, 102, 103, 104, 105, 106, and 107. Where, • Profile 1 is used to activate 2-wire SHDSL ports in the ATM mode. • Profile 100 is used to activate 4-wire SHDSL ports in the ATM mode. • Profile 101 is used to activate 6-wire SHDSL ports in the ATM mode. • Profile 102 is used to activate 8-wire SHDSL ports in the ATM mode. • Profile 103 is used to activate the SHDSL port bound to the EFM. • Profile 104 is used to activate 4-wire SHDSL ports in the TDM mode, and the frame encapsulation format is E1. • Profile 105 is used to activate 4-wire SHDSL ports in the TDM mode, and the frame encapsulation format is V35. • Profile 106 is used to activate 2-wire SHDSL ports in the TDM mode, and the frame encapsulation format is E1. • Profile 107 is used to activate 2-wire SHDSL ports in the TDM mode, and the frame encapsulation format is V35.
SHDSL alarm profile	Profile ID: 1

Procedure

- Configure an SHDSL line profile.

  Run the **shdsl line-profile quickadd** command to quickly add an SHDSL line profile, or run the interactive **shdsl line-profile add** command to add an SHDSL line profile.

  Main parameters:

- **data path mode**: Indicates the data path mode. Configure the data path mode according to the actual application scenario of the line. Three modes, namely ATM, PTM, and TDM modes are supported.
- **rate**: indicates the line rate. During line activation, a proper rate between the preset maximum rate and minimum rate is determined through automatic negotiation according to the line condition and the profile configuration. The user rate can be restricted by this rate or the rate set in the traffic profile that is bound to the user. When both rates function, the lower rate is selected as the user rate.
- **transmission**: indicates the transmission mode. Set the transmission mode according to line conditions and actual planning. Three transmission modes are supported: annex A, annex L, and annex A&B.
- **snr-margin**: The larger the SNR margin, the better the line stability, and meanwhile the lower the physical connection rate of the line after activation. For common Internet access users, set the target SNR margin to 3; for users with higher priorities, set the target SNR margin to 5.

NOTE
When the board supports G.SHDSL.bis (including the extended standard annex F), the maximum rate can reach 5696 kbit/s.

- Configure an SHDSL alarm profile.

  Run the **shdsl alarm-profile quickadd** command to quickly add an SHDSL alarm profile, or run the interactive **shdsl line-profile add** command to add an SHDSL alarm profile.

----End

Example

To add SHDSL line profile 3 with the line rate of 4096 kbit/s, which is used to activate the 4-wire SHDSL port, do as follows:

```
huawei(config)#shdsl line-profile quickadd 3 line four-wire rate 4096
```

Assume that the loop attenuation threshold is 10 dB, SNR margin is 0 dB, ES threshold is 100s, SES threshold is 100s, CRC abnormality duration threshold is 10000, LOSWS threshold is 100s, UAS threshold is 100s. To quickly add SHDSL line alarm profile 3 with these parameters, do as follows:

```
huawei(config-if-shl-0/3)#shdsl alarm-profile quickadd 3 loop-attenuation 10
snr-margin
 0 es 100 ses 100 crc-anomaly 10000 losws 100 uas 100
```

8.4.2 Configuring SHDSL Line Bonding

To ensure longer access distance at the same access rate or higher access rate in the same access distance, configure SHDSL line bonding.

Prerequisites

- The port to be bound has no service flow.
- The port to be bound is in the activating or deactivated state.

NOTE
An xDSL port can be in any of the following states: activating, activated, deactivated, and loopback.

Procedure

In the global config mode, run the **interface shl** command to enter the SHDSL mode.

Step 1 Run the **port bind m-pair** command to configure the SHDSL M-pair bonding. Run the **port bind efm** command to configure the SHDSL EFM M-pair bonding.

NOTE

- Inter-chip bonding is not supported. On an SHDSL board, ports 0-3 share one chip, ports 4-7 share one chip, ports 8-11 share one chip, and ports 12-15 share one chip. The ports to be bonded must be activated at the same time and must use the same line profile.
- Different line profiles can be applied to the ports in an EFM bonding group. When one port goes offline, the status of the entire binding group remains unchanged.
- When the SHDSL board supports G.SHDSL.bis (including the extended standard annex F), 1-pair bonding, 2-pair bonding, 3-pair bonding, and 4-pair bonding are supported, corresponding to the maximum available bandwidth of 5696 x M (M is the pair number; M is 1, 2, 3, or 4.) kbit/s. When the SHDSL board supports only G.991.2 (version 1), 2-pair bonding and 4-pair bonding are supported.
- After ports are bonded, all operations must be performed on the primary port.
- To delete a bonding group, only the ID of the primary port can be input.

----End

Example

The board chipset is in the ATM mode. To quadruple the bandwidth of a single port on SHDSL board 0/4 through m-pair bonding, do as follows:

```
huawei(config)#interface shl 0/4
huawei(config-if-shl-0/4)#port bind m-pair 8-11
```

8.4.3 Configuring an SHDSL Port

An xDSL port can transmit services only when it is activated. This topic describes how to activate an SHDSL port and bind an SHDSL profile to the port.

Prerequisites

8.4.1 Configuring SHDSL Profiles has been completed based on the data plan.

Procedure

Run the **interface shl** command to enter the SHDSL mode.

Step 1 Run the **activate** command to activate an SHDSL port and bind an SHDSL line profile to the port.

Step 2 Run the **alarm-config** command to bind an alarm profile to the port.

----End

Example

To activate SHDSL port 0/3/0 and bind line profile 2 and alarm profile 2 to the port, do as follows:

```
huawei(config)#interface shl 0/3
huawei(config-if-shl-0/3)#deactivate 0
huawei(config-if-shl-0/3)#activate 0 2
huawei(config-if-shl-0/3)#alarm-config 0 2
```

9 ATM Cascading

About This Chapter

The MA5600T/MA5603T provides ATM ports for cascading traditional ATM DSLAMs on a live network.

9.1 Introduction

Definition

The ATM access is a feature by which the MA5600T/MA5603T provides ATM ports to subtend the traditional ATM DSLAMs in the current network.

Purpose

Currently, the IP MAN, instead of the ATM network, is mainly used. Original ATM networks gradually evolve to IP MANs. In the evolution from ATM networks to IP networks, carriers are gradually replacing ATM devices with IP devices. In the current network, however, there are still a large number of ATM devices, which are distributed at the ATM access layer and the ATM backbone layer. To protect the investment and the network stability of carriers, the MA5600T/MA5603T, a new generation IP-core DSLAM, provides ATM ports to subtend the traditional ATM DSLAMs.

Glossary

Table 9-1 Glossary of the ATM access feature

Glossary	Explanation
PWE3	Pseudo wire emulation edge-to-edge (PWE3) is an end-to-end technology for bearing Layer 2 services. It is a point-to-point L2VPN.

Acronyms and Abbreviations

Table 9-2 Acronyms and abbreviations of the ATM access feature

Acronym/Abbreviation	Full Spelling
ATM	Asynchronous Transfer Mode
CAR	Committed Access Rate
PWE3	Pseudo wire Emulation Edge-to-Edge
PVC	Permanent Virtual Channel
PVP	Permanent Virtual Path
VP	Virtual Path

9.2 Principle

Clock Feature of the AIUG Board

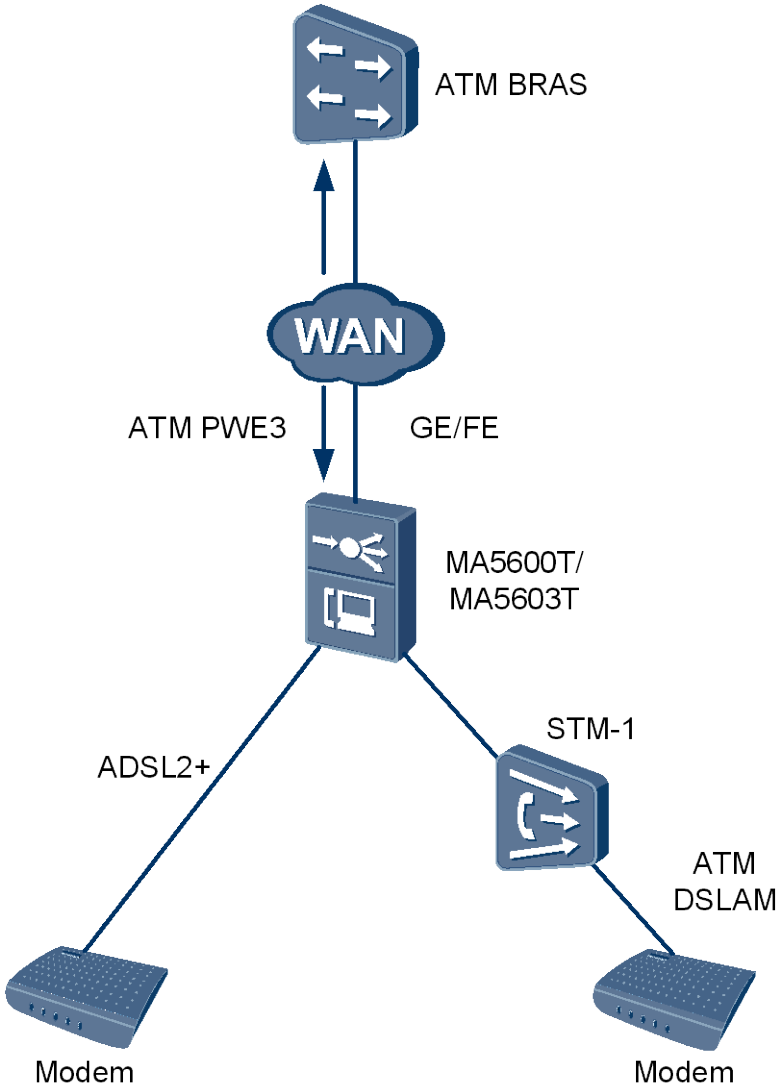
The AIUG board supports two modes of Tx clock: the system clock and the line clock.

The line-side clock of the AIUG board can be used as the clock source of the clock daughter board of the control board. At the same time, the system clock can be used as the line Tx clock of the AIUG board. When the control board does not have a clock daughter board, the system clock can be used as the line-side clock of the AIUG board.

ATM Access/Upstream Transmission Through Ethernet Ports

In the case of the ATM access, the upstream transmission through Ethernet ports is supported. The most common function of an ATM port is to convert the ATM cells from the ATM DSLAM into Ethernet packets, and then to send the Ethernet packets to the upper-layer Ethernet MAN through the upstream interface of the IP DSLAM. Figure 9-1 illustrates the principle of ATM access/upstream transmission through Ethernet ports.

Figure 9-1 Principles of ATM access/upstream transmission through Ethernet ports



- Upstream direction (from the ATM DSLAM to the IP DSLAM)
  a. Restore the ATM frames from the ATM DSLAM to ATM cells.
  b. Assemble ATM cells to ALL5 frames.
  c. Restore AAL5 frames to Ethernet frames.
  d. Add the corresponding VLAN tag in the Ethernet frame header and send the Ethernet frame to the Ethernet MAN through the upstream interface.
- Downstream direction (from the IP DSLAM to the ATM DSLAM)
  a. The IP DSLAM receives Ethernet packets from the Ethernet MAN and encapsulates them to AAL5 frames.
  b. The IP DSLAM segments AAL5 frames as single cells.
  c. The IP DSLAM encapsulates cells to the frames of the corresponding ATM interface (for example, an STM-1 port) and sends the frames to the ATM DSLAM through the ATM interface (for example, an STM-1 port).

9.3 Configuring the ATM-DSLAM Access Service

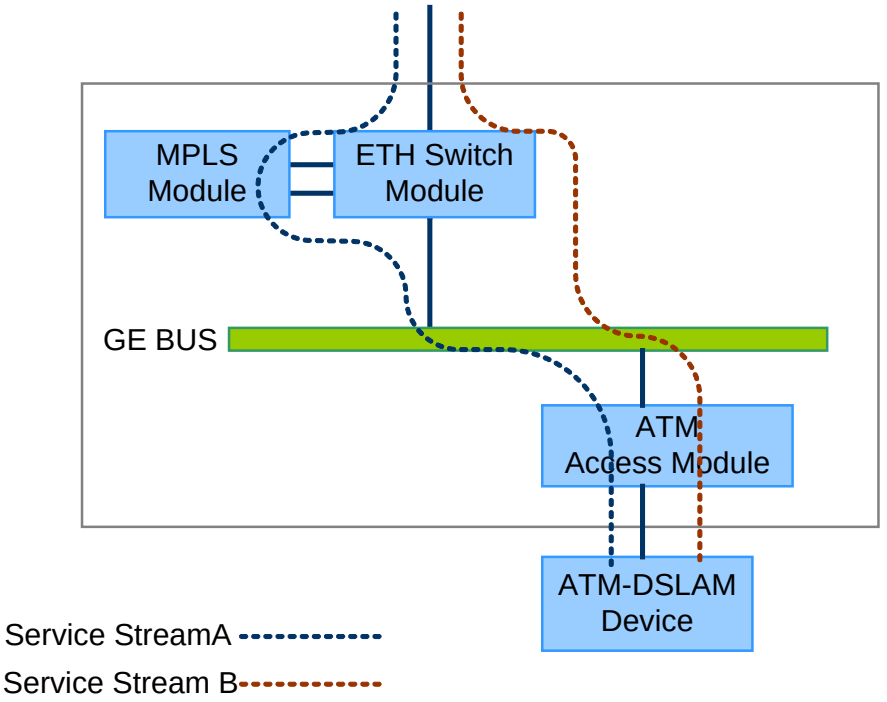
ATM-DSLAM access means that the MA5600T/MA5603T/MA5608T provides the ATM interface (for example, STM-1) for the subtending of earlier ATM-DSLAMs.

Context

In the evolution from ATM networks to IP networks, carriers will replace their ATM-DSLAM network devices in the access layer with IP network devices. In this evolution, a large number of ATM network devices still exist in the network for a long time. The MA5600T/MA5603T/MA5608T provides ATM ports for lower level ATM network devices to access the network.

The MA5600T/MA5603T/MA5608T provides four ATM optical ports (STM-1) through the AIUG board for connecting to the ATM-DSLAM, and also provides the common Ethernet upstream or MPLS upstream service, as shown in Figure 9-2.

Figure 9-2 ATM-DSLAM access



The MA5600T/MA5603T/MA5608T can provide two upstream transmission modes: direct Ethernet upstream transmission mode and MPLS upstream transmission mode.

- Directly Ethernet upstream transmission mode: Traffic stream B of the ATM-DSLAM is directly transmitted upstream to the upper-layer IP network through the Ethernet switching module of the SCU board. This mode is applicable to the common Internet access service.
- MPLS upstream transmission mode: The MA5600T/MA5603T/MA5608T functions as a provider edge (PE), transmitting and services of the subtended ATM-DSLAM through the upstream port to the MPLS network. This mode is applicable to the private line service. According to actual requirements, the data on the upstream port can be encapsulated in the ATM PWE3 mode or the ETH PWE3 mode.

- ATM PWE3: The MA5600T/MA5603T/MA5608T creates a transparent transmission channel for private line users. After encapsulated in the ATM PWE3 mode, the data is transmitted upstream to the MPLS core network. After reaching the peer device, the data is decapsulated, and the ATM cells are transmitted downstream to peer users. This encapsulation mode is applicable to the scenario where the ATM-DSLAM needs to communicate with the peer ATM-DSLAM or peer ATM BRAS over the MPLS network.
- ETH PWE3: The MA5600T/MA5603T/MA5608T creates a transparent transmission channel for users. After encapsulated in the ETH PWE3 mode, the data is transmitted upstream to the MPLS core network. After reaching the peer device, the data is decapsulated. This encapsulation mode is applicable to the scenario where xPoA private line users perform authentication and packet forwarding over the MPLS network.

NOTE
For xPoA users, the xPoA to xPoE protocol conversion should be configured.

9.4 Reference Standards and Protocols

The following lists the reference standards and protocols of this feature:

- ITU-T I.363.5, AAL5 Service Adaptation Protocol
- ITU-T I.361, B-ISDN ATM layer specification

10 MPLS

About This Chapter

Multiprotocol Label Switching (MPLS) was introduced to improve the forwarding speed. However, because of its excellent performance in traffic engineering (TE) and virtual private network (VPN), which are the two critical technologies, MPLS is becoming an important standard for extending the IP network.

10.1 Overview

Multi-protocol Label Switching (MPLS) is between the data link layer and the network layer in the TCP/IP protocol stack. The label in a short fixed length is used to encapsulate IP packets. On the data plane, fast label forwarding is implemented. On the control plane, MPLS can meet the requirements on the network from various new applications with the help of the powerful and flexible routing functions of the IP network.

The MPLS feature includes the following sub features:

- Basic MPLS functions

  Basic MPLS functions provide a basis for other MPLS sub features. MPLS, which is not restricted by any specific link layer protocol, can use any Layer 2 medium to transmit network packets. This shows that MPLS is not a service or application, but a tunnel technology. This technology can both support multiple higher-layer protocols and services, and ensure the security of information transmission to a certain extent.
- MPLS RSVP-TE

  To deploy engineered traffic on a large-scale backbone network, a simple solution with good expansibility must be adopted. MPLS, as a stacking model, can easily establish a virtual topology over a physical network and map traffic to this topology. Therefore, a technology that integrates MPLS with traffic engineering, namely, MPLS-TE is generated.
- MPLS OAM

  MPLS, as the key bearer technology for the extensible network-generation network, provides multiple services with QoS guarantee. In addition, MPLS introduces a unique network layer and therefore the faults caused by this new network layer may occur. Therefore, an MPLS network must have the OAM capability.

The MPLS feature supports the following functions:

- Functioning as a P device
- Capability of 100 pps for processing LDP and RSVP packets when functioning as a P device
- MPLS label switching
- Penultimate hop popping (PHP)
- Query of LSP packet statistics by label

10.2 Reference Standards and Protocols

The following lists the reference standards and protocols of this feature:

1. PWE3
   - RFC3985: Pseudo Wire Emulation Edge-to-Edge (PWE3) Architecture
   - RFC4447: Pseudowire Setup and Maintenance Using the Label Distribution Protocol (LDP)
   - RFC3916: Requirements for Pseudo-Wire Emulation Edge-to-Edge (PWE3)
   - RFC4446: IANA Allocations for Pseudowire Edge to Edge Emulation (PWE3)
   - RFC4717: Encapsulation Methods for Transport of Asynchronous Transfer Mode (ATM) over MPLS Networks
   - RFC4448: Encapsulation Methods for Transport of Ethernet over MPLS Networks
   - RFC5085: Pseudowire Virtual Circuit Connectivity Verification (VCCV): A Control Channel for Pseudowires
   - RFC4553: Structure-Agnostic Time Division Multiplexing (TDM) over Packet (SAToP)
   - RFC5462: Multiprotocol Label Switching (MPLS) Label Stack Entry: EXP Field Renamed to Traffic Class Field
   - RFC4385: Pseudowire Emulation Edge-to-Edge (PWE3) Control Word for Use over an MPLS PSN
   - draft-ietf-pwe3-redundancy-bit-00
2. RSVP
   - RFC2205: Resource ReSerVation Protocol (RSVP) -- Version 1 Functional Specification
   - RFC3209: RSVP-TE: Extensions to RSVP for LSP Tunnels
   - RFC2210: The Use of RSVP with IETF Integrated Services
   - RFC2961: RSVP Refresh Overhead Reduction Extensions
   - RFC3270: Multi-Protocol Label Switching (MPLS) Support of Differentiated Services
   - RFC4090: Fast Reroute Extensions to RSVP-TE for LSP Tunnels
3. LDP
   - RFC3031: Multiprotocol Label Switching Architecture
   - RFC5036: LDP Specification
   - RFC3215: LDP State Machine
   - RFC3478: Graceful Restart Mechanism for Label Distribution Protocol

- RFC3815: Definitions of Managed Objects for the Multiprotocol Label Switching (MPLS), Label Distribution Protocol (LDP)

4. MPLS
   - draft-ietf-mpls-lsp-ping-version-06
   - RFC4379: Detecting Multi-Protocol Label Switched (MPLS) Data Plane Failures
   - RFC3032: MPLS Label Stack Encoding
   - RFC3469: Framework for Multi-Protocol Label Switching (MPLS)-based Recovery
   - RFC3812: Multiprotocol Label Switching (MPLS) Traffic Engineering (TE) Management Information Base (MIB)
   - RFC3813: Multiprotocol Label Switching (MPLS) Label Switching Router (LSR) Management Information Base (MIB)
   - RFC3814: Multiprotocol Label Switching (MPLS) Forwarding Equivalence Class To Next Hop Label Forwarding Entry (FEC-To-NHLFE) Management Information Base (MIB)
   - Y.1710: Requirements for OAM functionality for MPLS networks
   - Y.1711: OAM mechanisms for MPLS networks
   - Y.1720: Protection switching for MPLS networks

10.3 MPLS

Multiprotocol Label Switching (MPLS) was introduced to improve the forwarding speed. However, because of its excellent performance in traffic engineering (TE) and virtual private network (VPN), which are the two critical technologies, MPLS is becoming an important standard for extending the IP network. This topic provides the introduction, availability, principle, and reference of the MPLS feature.

10.3.1 Introduction

Definition

Basic MPLS features mainly refer to the MPLS Label Distribution Protocol (LDP) and LSP management function.

The LDP protocol is a standard MPLS label distribution protocol defined by the IETF. LDP, which is mainly used to allocate labels for the negotiation between LSRs to set up label switching paths (LSPs), regulates various types of information for the label distribution process, and the related processing. The LSRs form an LSP that crosses the entire MPLS domain according to the local forwarding table, which correlates in the label, network hop node, and out label of each specific FEC.

With the LSP management function, the MA5600T/MA5603T/MA5608T can manage and maintain the LSPs generated by various LDPs and can issue the hardware forwarding module.

Purpose

MPLS is initially put forth to improve the forwarding speed of routers. Compared with the traditional IP routing mode, during data forwarding, MPLS analyzes the IP packet header only on the edge of the network, but does not analyzes the IP packet header at each hop. This saves the processing time.

With the development of the ASIC technology, the route search speed is not a bottleneck for network development. Thus, MPLS has not obvious advantages in forwarding speed. MPLS, however, is widely applied to the virtual private network (VPN), traffic engineering, and quality of service (QoS) due to its characteristics of supporting multi-layer labels and connected-oriented forwarding plane. Therefore, MPLS becomes an increasingly important standard for expanding the scale of the IP network.

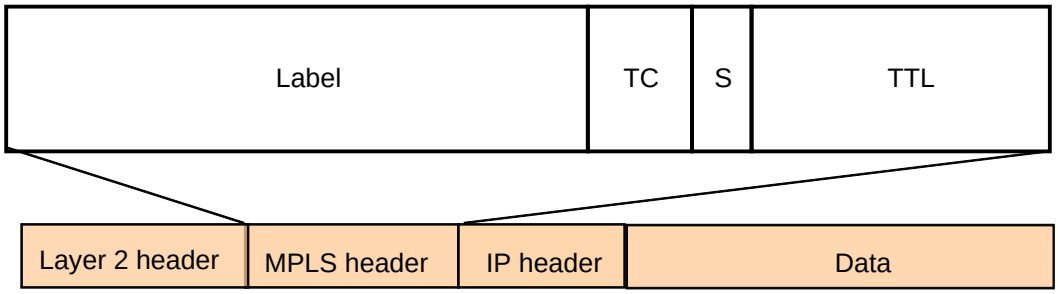
10.3.2 Principle

Multiprotocol label switching (MPLS) was introduced to improve the forwarding speed. However, because of its superb performance in traffic engineering (TE) and virtual private network (VPN), which are the two critical technologies in the current IP network, MPLS has become an important standard for extending the IP network.

IP technologies are connectionless at both the forwarding plane and control plane while ATM technologies are connection-oriented at the two planes. The MPLS technology combines the advantages of IP and ATM technologies and achieves a connectionless control plane and a connection-oriented forwarding plane. Such a combination provides for flexible IP routing and convenient Layer 2 switching as well as expanded ATM service variety.

Figure 10-1 shows the MPLS packet format.

Figure 10-1 MPLS packet format



- Label: a 20-bit label value field, used as the forwarding pointer.
- TC: short for traffic class, a 3-bit field for QoS (note that this field was named EXP and is renamed TC in RFC5462).
- S: a 1-bit bottom of stack field. This bit set to 1 indicates the bottom label in the label stack.
- TTL: short for time to live, an 8-bit field, similar to the TTL field in an IP header.

Basic MPLS Concepts

- Forwarding equivalence class (FEC)

  An FEC refers to a group of data streams which are forwarded in the same manner. These data streams are forwarded by the LSR in the same manner. Theoretically, FECs can be classified according to the IP address, service type, or QoS. For example, in the conventional IP forwarding by using the maximum matching algorithm, all the packets to the same route belong to an FEC. Currently, FECs are generally classified based on the address. The MA5600T/MA5603T/MA5608T supports only address-based FECs.

- Label

  A label is a short fixed length physically contiguous identifier which is used to identify an FEC, usually of local significance. In certain conditions, for example, when load

sharing is required, one FEC may map multiple labels. On one device, however, one label can represent only one FEC.

Label encapsulation is performed between the link layer and the network layer. Therefore, label can be supported by any link layer.

- Penultimate hop popping

  On the last hop node, the label no longer has any function. In this case, the label stack may be popped at the penultimate LSR of the LSP, rather than at the LSP Egress, to reduce the load of the last hop LSR. The last hop LSR directly forwards IP packets or next-layer labels, which are configured at the egress by the PHP.

- Label switching router (LSR)

  An LSR, also called an MPLS node, is a network device which is capable of exchanging and forwarding MPLS labels. LSRs are the basic elements in an MPLS network. All LSRs support the MPLS protocol.

- Label edge router (LER)

  An LSR on the edge of the MPLS domain is called the LER. If an LSR has a neighbor node that does not run the MPLS protocol, the LSR is an LER.

  The LER is responsible for classifying the packets that enter the MPLS domain to FECs and adding labels to these FECs for forwarding in the MPLS domain. When the packets leave the MPLS domain, the FECs pop up the labels, resume the original packets, and then are forwarded accordingly.

- Label switched path (LSP)

  The path that a packet in a particular FEC traverses in an MPLS network is called the LSP.

  The LSP, similar to the ATM virtual circuit in function, is a unidirectional path from the ingress to the egress.

- Label distribution protocol (LDP)

  LDP, also called the signaling protocol, is the MPLS control protocol. LDP is responsible for series of operations such as FEC classification, label distribution, and LSP establishment and maintenance.

  MPLS can use multiple label distribution protocols, such as the Label Distribution Protocol (LDP) and Resource Reservation Protocol Traffic Engineering (RSVP-TE).

  - LDP is a standard MPLS label distribution protocol defined by the IETF. LDP is responsible for FEC classification, label distribution, and LSP establishment and maintenance.
  - RSVP-TE is an extension to RSVP and provides high QoS and TE capability for users by establishing TE LSPs.

- Label distribution mode

  In an MPLS system, the downstream LSR determines the label to be advertised to a specific FEC, and then notifies the upstream LSR. That is, the label is specified by the downstream LSR, and is advertised from the downstream LSR to the upstream LSR.

  The label advertisement modes on the upstream and downstream LSRs with label advertisement adjacencies must be the same. Otherwise, the LSP cannot be set up.

  The two label advertisement modes are as follows:

  - Downstream unsolicited (DU) mode

    In the DU mode, the LSR allocates labels to a specific FEC without asking for the label request message from upstream LSRs.

  - Downstream on demand mode

In the DoD mode, the LSR allocates labels to a specific FEC only after obtaining the label request message from upstream LSRs.

NOTE

When a downstream LSR feeds back the label mapping information is determined by the label control mode used by the LSR.

- When an LSR supports the ordered label control mode, it sends the label mapping information to the upstream LSR only when it receives the label mapping message returned by the downstream LSR, or when it is the egress node of the FEC.
- When an LSR supports the independent label distribution control mode, it sends the label mapping message to the upstream LSR regardless of whether it receives the label mapping message returned by the downstream LSR.

- Label distribution control mode

  The label distribution control mode is the mode used by the LSR to allocate labels during the establishment of LSPs.

  The two label distribution control modes are as follows:

  - Independent label distribution control mode

    In the independent label distribution control mode, the local LSR can independently allocate a label to an FEC and binds the label to the FEC, and notify the upstream LSR of the label, without waiting for the label from the upstream LSR.

  - Ordered label control mode

    In the ordered label control mode, the LSR can send the label mapping message of an FEC to the upstream LSR only when the LSR has the label mapping message of the next hop of the FEC, or when the LSR is the egress node of the FEC.

- Label retention mode

  The label retention mode is the mode adopted by the LSR to process the received label mapping messages that are not in use temporarily.

  The two label retention modes are as follows:

  - Liberal retention mode

    If an LSR supports the liberal retention mode, it maintains the label mapping received from the neighbor LSR regardless of whether the neighbor LSR is its own next hop.

    When the next hop neighbor changes due to the change of network topology, the LSR that supports the liberal retention mode can use the label sent from the non-next-hop neighbor to set up LSPs quickly. This, however, requires more memory and label space.

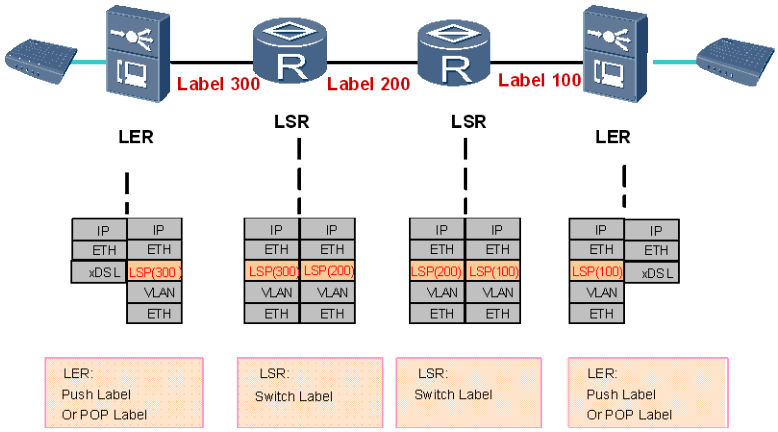
  - Conservative retention mode

    If an LSR supports the conservative retention mode, it maintains the label mapping received from the neighbor LSR only when the neighbor LSR is its next hop.

    When the next hop neighbor changes due to the change of network topology, the LSR that supports the conservative retention mode can save memory and label space because the LSR maintains only the label from the next hop neighbor. The re-establishment of LSPs, however, lasts a long time.

Figure 10-2 shows the protocol stack model for label distribution.

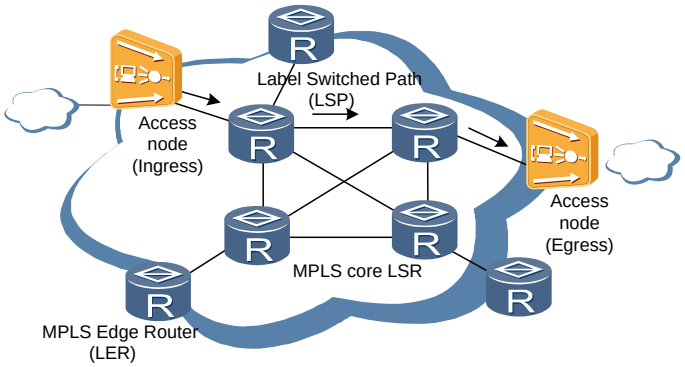
Figure 10-2 Protocol stack model for label distribution



Working principle of the MPLS feature

Figure 10-3 shows the working principle of the MPLS feature

Figure 10-3 MPLS network structure

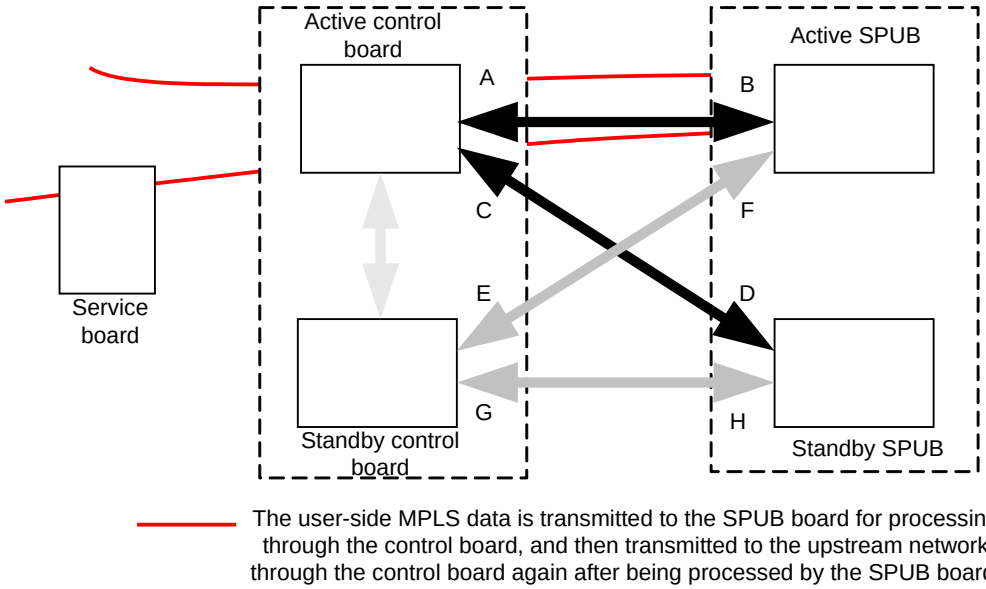


1. First, enable MPLS and LDP on each router on the network, and enable LDP on the interconnected interfaces.
2. Consequently, LDP automatically sets up an LDP session between any two routers. The LDP packets are carried on this session.
3. LDP works with the traditional routing protocol such as OSPF and RIP to set LSPs in each LSR for the FEC with service requirements.
4. LDP does not need to be enabled for the establishment of static LSPs. Configure the FEC, and inbound and outbound labels on each MPLS router that the static LSP travels.

MPLS Active and Standby Protection

The MA5600T/MA5603T/MA5608T implements active and standby protection for the MPLS service through the active and standby MPLS service boards (SPUBs). Figure 10-4 shows the working principle of active and standby protection for the MPLS service.

Figure 10-4 Working principle of active and standby protection for the MPLS service



Port B of the two internal 10GE ports on the active SPUB board is connected to port A on the active control board. Ports A and B are used to receive and transmit the network-side and user-side packets. The other port (port F) is connected to port E on the standby control board.

Port D of the two internal 10GE ports on the standby SPUB board is connected to port C on the active control board. Ports C and D are used to receive and transmit the network-side and user-side packets. The other port (port H) is connected to port G on the standby control board.

Therefore, after the active and standby SPUB boards form a protection group, the system automatically switches the MPLS services to the standby SPUB board when the active SPUB board fails, thereby implementing active and standby protection for the MPLS services.

LDP GR

The GR is a key technology for implementing the high availability (HA). The GR protocol collects the information about the protocol control plane from neighbors or remote peers but does not learn about the information about the control plane through the handshake and exchange of the protocol.

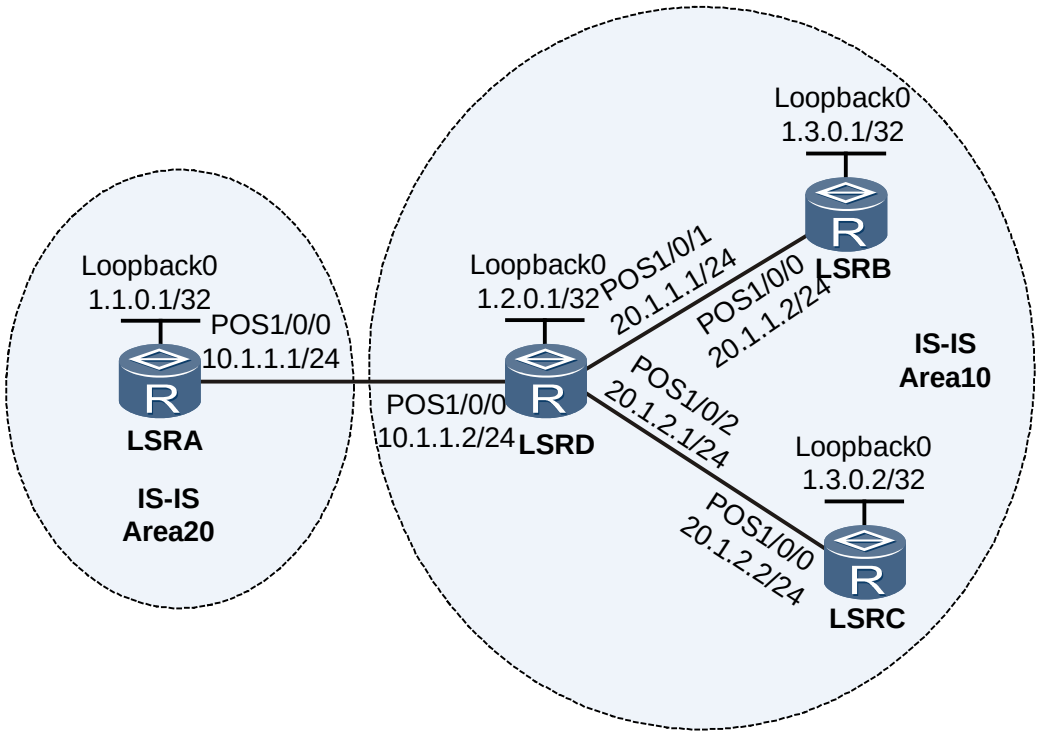
The LDP GR function ensures normal forwarding of the MPLS service during the active/standby switchover or upgrade of the system. In addition, the LDP GR function resumes the LDP session and completes the LSP establishment after the active/standby switchover or upgrade of the system

NOTE

In actual application, to prevent services from being affected by the active control board failure, configure the system-level GR in the environment where both active and standby control boards are configured.

LDP Extension for Inter-Area LSP

Figure 10-5 Networking topology of LDP Extension for Inter-Area LSP



As shown in Figure 10-5, there are two IGP areas, Area 10 and Area 20.

In the routing table of LSRD at the edge of Area 10, there are two host routes to LSRB and LSRC. Generally, to prevent a large number of routes from occupying too many resources, on LSRD, you can use IS-IS to aggregate the two routes to one route 1.3.0.0/24 and send this route to Area 20. Consequently, there is only one aggregated route (1.3.0.0/24) but not 32-bit host routes in the routing table of LSRA. By default, when establishing LSPs, LDP searches the routing table for the route that exactly matches the forwarding equivalence class (FEC) in the received Label Mapping message. Table 10-1 shows routing entry information of LSRA and routing information carried in FEC in the situation as shown in Figure 10-5.

Table 10-1 Routing entry information of LSRA and routing information carried in FEC

Routing entry information of LSRA	FEC
1.3.0.0/24	1.3.0.1/32
	1.3.0.2/32

LDP establishes liberal LSPs rather than inter-area LDP LSPs for aggregated routes. In this situation, LDP cannot provide required backbone network tunnels for VPN services.

Therefore, in the situation as shown in Figure 10-5, you need to configure LDP to search for routes according to the longest match rule to establish LSPs. There is already an aggregated

route 1.3.0.0/24 in the routing table of LSRA. When LSRA receives a Label Mapping message (such as the carried FEC is 1.3.0.1/32) from Area 10, LSRA searches for a route according to the longest match rule defined in RFC 5283. Then, LSRA finds information about the aggregated route 1.3.0.0/24, and uses the outbound interface and next hop of this route as those of the route 1.3.0.1/32. In this manner, LDP can establish inter-area LDP LSPs.

10.4 MPLS RSVP-TE

MPLS RSVP-TE is a technology which integrates TE and the MPLS superimposed model. It provides high quality of service (QoS) and TE capability for users by establishing LSPs based on TE. This topic provides introduction to this feature and describes the principle and reference documents of this feature.

10.4.1 Introduction

Definition

MPLS RSVP-TE is a technology that integrates TE with the MPLS technology. MPLS RSVP-TE establishes label switched path (LSP) tunnels along specified paths for resource reservation, enables network traffic to avoid the node where congestion occurs to balance network traffic.

To establish constraint-based LSPs in MPLS TE, RSVP is extended. The extended RSVP signaling protocol is called the RSVP-TE signaling protocol.

Purpose

To deploy engineered traffic on a large-scale backbone network, a simple solution with good expansibility must be adopted. MPLS, as a stacking model, can easily establish a virtual topology over a physical network and map traffic to this topology.

MPLS TE establishes the LSP tunnel along a specified path through RSVP-TE and reserves resources. Thus, carriers can accurately control the path that traffic traverses to avoid the node where congestion occurs. This solves the problem that certain paths are overloaded and other paths are idle, utilizing the current bandwidth resources sufficiently. At the same time, MPLS TE can reserve resources during the establishment of LSP tunnels to ensure the QoS.

To ensure continuity of services, MPLS TE also introduces route backup to implement quick switching in case of link failure.

10.4.2 Principle

Basic MPLS RSVP-TE Concepts

- CR-LSP

  An LSP that is established based on certain constraints is called a constraint-based routed label switched path (CR-LSP). Different from a common LSP, the establishment of a CR-LSP depends on the routing information. In addition, some conditions must be met, for example, the specified bandwidth, the fixed route, and QoS parameters.

  CR-LSPs can be classified into the following two categories:

  - Static CR-LSP

The forwarding information and resources information about a static CR-LSP are configured manually and the signaling protocol and route calculation are not involved. Less resource is occupied because the MPLS control packets do not need to be exchanged. The static CR-LSP, however, is seldom applied because it cannot dynamically adjust according to the topology change of the network.

- Dynamic CR-LSP

  A dynamic CR-LSP is established and maintained through the signaling mechanism, and route calculation is required.

- RSVP

  Resource Reservation Protocol (RSVP) is designed for the integrated service model and is used to reserve resources on each node on a path. RSVP works on the transmission layer, but does not participate in the transmission of application data. RSVP, similar to ICMP, is a network control protocol.

- RSVP-TE

  To establish the CR-LSP, RSVP is extended. The extended RSVP signaling protocol is called the RSVP-TE signaling protocol.

- Explicit route

  A CR-LSP that is established along a specified path is called an explicit route. The two types of explicit route are as follows:

  - Strict explicit route

    On a strict explicit route, the next hop node must be directly connected to its preceding hop node. The route of the LSP can be precisely controlled by using the strict explicit route.

  - Loose explicit route

    The path between a loose node and its preceding node MAY include other network nodes that are not part of the strict node or its preceding abstract node.

NOTE

The MPLS TE signaling can carry the strict or loose attributes of an explicit path, and establish a CR-LSP along a specified path.

Composition of MPLS RSVP-TE

The following four components are necessary to the MPLS TE function:

- Information advertisement component

  In addition to the topology information about the network, TE also needs to know the load information about the network. Therefore, MPLS TE introduces the information advertisement component, that is, MPLS TE maintains the link attribute and topology attribute of the network on each node through IGP extensions to form the TE database (TEDB). The path that meets all types of constraints can be calculated by using the TEDB. The extended OSPF protocol adds certain TE-related attributes such as link bandwidth and color to the link connection status, where the maximum reservable bandwidth and unreserved bandwidth for the link with each priority are the most important.

- Route selection component

  After the information advertisement component forms the TEDB, the path that the LSP tunnel passes can be specified on each ingress node. This explicit path can be a strict or loose explicit path. In addition, the restraints such as the bandwidth can be specified.

  The route selection component calculates the path that meets the specified constraints by using the data in the TEDB through the constraint shortest path first (CSPF) algorithm.

- Signaling component

  After the shortest path from the ingress to the egress of the LSP is obtained, the TE tunnel, which is used to forward the traffic that enters the ingress of the LSP, needs to be established. This process is implemented by the signaling component.

  The MA5600T/MA5603T/MA5608T supports establishment of LSP tunnels through RSVP. The RSVP signaling can carry the constraint parameters such as the bandwidth of the LSP, certain explicit routes, and color.

  An LSP can also be established without the signaling protocol. That is, an LSP can be established through allocating labels manually hop by hop. An LSP established in this mode is called a static CR-LSP.

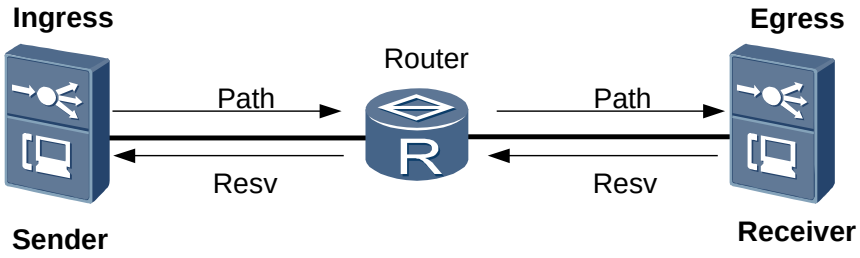
- Packet forwarding component

  The packet forwarding component of MPLS RSVP-TE is based on the label, that is, it forwards packets along the existing LSPs through labels. The defects of the IGP routing protocol can be avoided because the path of an LSP tunnel can be specified.

Process of TE LSP Tunnel Establishment

The LSP established through RSVP-TE has the resource reservation capability, and certain resources of the LSR on the LSP can be allocated to the LSP. Thus, the services transmitted on the LSP can be guaranteed. Figure 10-6 shows the process of TE LSP tunnel establishment.

Figure 10-6 Process of TE LSP tunnel establishment



The process of TE LSP tunnel establishment is summarized as follows:

1. The ingress LSR generates the Path message and transmits it to the egress LSR.
2. After the egress LSR receives the Path message, the egress LSR generates the Resv message and transmits it to the ingress LSR. At the same time, the LSRs on the LSP reserves resources for the LSP through the Resv message.
3. When the ingress LSR receives the Resv message, it indicates that the LSP is successfully established.

RSVP-TE GR

RSVP-TE graceful restart (GR) is a status recovery mechanism of RSVP-TE. When the control plane performs active/standby switchover, RSVP-TE GR can ensure the continuity of data transmission on the forwarding plane. At the same time, neighbor nodes help the GR node to recover in time.

RSVP-TE GR is based on the Hello mechanism of RSVP. The recovery of the local status depends on the upstream Path message or the downstream Recovery Path message.

RSVP GR has the following features: Shortening the information recovery of the control plane; reducing changes of temporary routes; ensuring the continuity of service forwarding on the forwarding plane.

10.5 MPLS OAM

MPLS OAM checks if an LSP is in the normal state through a mechanism, and reports the alarm information if the LSP fails. This topic provides introduction to this feature and describes the principle and reference documents of this feature.

10.5.1 Introduction

Definition

Operation Administration & Maintenance (OAM) has the following features:

- Simplifying network operations
- Checking the network performance anytime
- Reducing OPEX of the network

Deployment of an effective OAM mechanism is crucial to the running of the network, especially to the network with certain QoS requirements, namely, certain performance and usability requirements.

MPLS, as the key bearer technology for the extensible network generation network, provides multiple services with QoS guarantee. In addition, MPLS introduces a unique network layer and therefore there will be faults that are only relevant to this new network layer. Therefore, an MPLS network must have the OAM capability.

MPLS OAM provides both detection tools and mature protection switching mechanisms. In this way, MPLS can perform switching when a fault occurs on the MPLS layer. This minimizes the loss of data.

Purpose

The MPLS OAM functions are as follows:

- Fault detection: Requirement-based query and continuous detection are provided to learn about anytime whether faults exist on the monitored LSP.
- Protection switching: After a fault occurs, it can be detected, analyzed, and located, and an alarm will be reported. In addition, the corresponding measures can be taken according to the fault type.

10.5.2 Principle

Background Knowledge for MPLS OAM

1. MPLS OAM packets are classified as follows:
   - Connectivity detection (CD) packets. The two types of CD packets are as follows:
     - Connectivity verification (CV)
     - Fast failure detection (FFD)

- Forward defect indication (FDI)
- Backward defect indication (BDI)

MPLS OAM is implemented by periodically transmitting detection packets CV or FFD over the detected LSPs.

2. Basic detection process

   MPLS OAM is implemented by periodically transmitting detection packets CV and FFD over the detected LSPs.

   - To detect the source by using the CV packet, a sliding window in the width of 3s is set on the source and the LSP status is checked by using the VC packet received in the sliding window.
   - To detect the source by using the FFD packet, a sliding window in the width of three times of FFD transmit interval is set on the source and the LSP status is checked by using the FFD packet received in the sliding window.

3. CV and FFD

   The FFD and CV detection packets are mutually exclusive. That is, only the FFD or CV detection packets can be applied to one LSP at a time.

4. Backward path

   BDI packets are transmitted through the backward path. The ingress of a backward path is the egress of the detected LSP, and the egress of the backward path is the ingress of the detected LSP. That is, each forward LSP has a backward path.

5. Protection switching (PS)

   When a fault occurs on the network, currently MPLS OAM provides the PS, a type of end to end tunnel protection technology, to recover the interrupted services.

   The PS uses one tunnel to protect another tunnel. There is no relation among the attributes of each tunnel in the protect group. For example, the protection tunnel with 10 Mbit/s bandwidth can protect a master tunnel with a requirement for 100 Mbit/s bandwidth.

MPLS OAM Detection Function

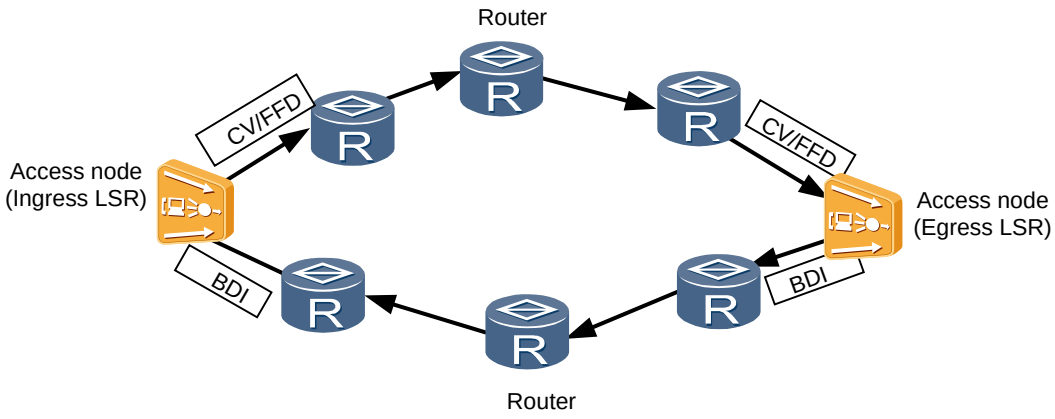
The basic process for MPLS OAM to detect the connectivity of a single LSP is as follows:

- The source transmits the CV/FFD packets to the destination through the detected LSP.
- The destination checks the correctness of the type and frequency information carried in the received detection packets and measures the number of correct and errored packets that are received within the detection period to monitor the connectivity of the LSP in real time.
- When the LSP fails, the destination detects the defect quickly and analyzes the defect type.

Bind a backward LSP to the detected LSP when configuring the OAM function for the detected LSP. A backward path is an LSP that has the opposite source and destination of the detected LSP, or a non-MPLS path that can be connected to the source and destination of the detected LSP.

After the destination detects a defect, the destination transmits the BDI packets that carry the defect information to the source through the backward path. The source learns about the status of the defect, and triggers the corresponding protection switching when the protect group is correctly configured. Figure 10-7 shows the MPLS OAM CD.

Figure 10-7 MPLS OAM CD



Working Modes of the MPLS OAM Protection Switching

The MPLS OAM protection switching aims at the entire LSP instead of one section or one node on the LSP. The route and bandwidth of the standby LSP for a specified active LSP are reserved.

Therefore, the protection switching is a thorough-assignment protection mechanism. To ensure that protection switching can be implemented effectively in all the possible cases that the active LSP fails, the standby LSP needs to use a physical path totally different from that of the active LSP.

The working mode of MPLS OAM protection switching is 1:1 protection mode. In this mode, each active LSP has a standby LSP.

- In normal conditions, data is transmitted through the active LSP and no traffic is transmitted through the standby LSP.
- When the destination detects a failure on the active LSP through the detection mechanism, the destination switches to the standby LSP, and then transmits the BDI packet to the source through the backward path, instructing the ingress to switch the traffic on the active LSP to the standby LSP. Thus, 1:1 protection switching is implemented.

10.6 MPLS TE Reliability

MPLS TE tunnels that transmit mission-critical services require high reliability. Access node supports the following network-level reliability.

- RSVP-TE FRR
- TE tunnel protection group
- CR-LSP backup

10.6.1 RSVP-TE FRR

RSVP-TE FRR is also called MPLS fast reroute. RSVP literally means the resource reservation protocol, TE means traffic engineering, and FRR means fast reroute.

Introduction

Definition

The RSVP-TE FRR technology is applied to the MPLS TE network for implementing partial network protection. Specifically, when a certain link or node in the network fails, the LSP configured with FRR can automatically switch the data to the protection link.

To ensure the reliability of the MPLS network, the MPLS FRR technology is combined with the MPLS TE technology to provide LSPs with fast switching. In the MPLS FRR, a local backup path is created beforehand to protect the LSP from the impact of the link or node failure. When a failure occurs, the device that detects the failure can quickly switch the service from the faulty link to the backup path, thus reducing data loss.

Purpose

Quick response and prompt switching are the features of MPLS FRR. Such features ensure the smooth switching of service data and prevent service interruption. In addition, the head node of the LSP will look for a new path for establishing a new LSP and will switch the service to the new LSP. Before the new LSP is set up, the service data is forwarded through the protection path.

Principle

MPLS TE FRR

The basic principle of MPLS TE FRR is to protect one or more LSPs by using an LSP that is created beforehand. The LSP that is created beforehand is called the FRR LSP (bypass LSP), and the LSP that is protected is called a primary LSP. The purpose of MPLS TE FRR is to bypass the faulty link or node through the bypass LSP to protect the primary LSP. Creating the bypass LSP and primary LSP requires the participation of all the components of the MPLS TE system.

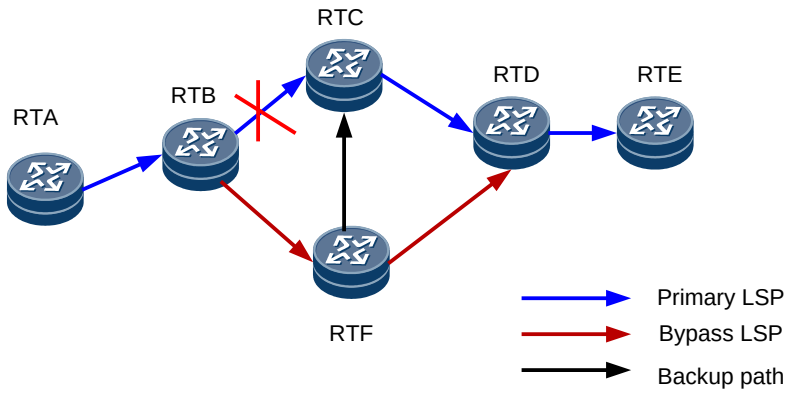
MPLS TE FRR is implemented based on RSVP TE and complies with RFC4090.

MPLS TE FRR can be implemented in the following two modes:

- Detour mode: This mode is also called the one-to-one backup mode. In this mode, one protection path is created to provide protection for each LSP. This protection path is called the detour LSP.
- Bypass mode: This mode is also called the facility backup mode. In this mode, one protection path provides protection for multiple LSPs. This protection path is called the bypass LSP.

The detour mode provides protection for each LSP, thus requiring more overheads. In the actual application, the bypass mode is more widely used. The MA5600T/MA5603T/MA5608T adopts the bypass mode. The following content of this topic mainly deals with the bypass mode. Figure 10-8 illustrates the FRR function implemented in the bypass mode.

Figure 10-8 FRR in the bypass mode

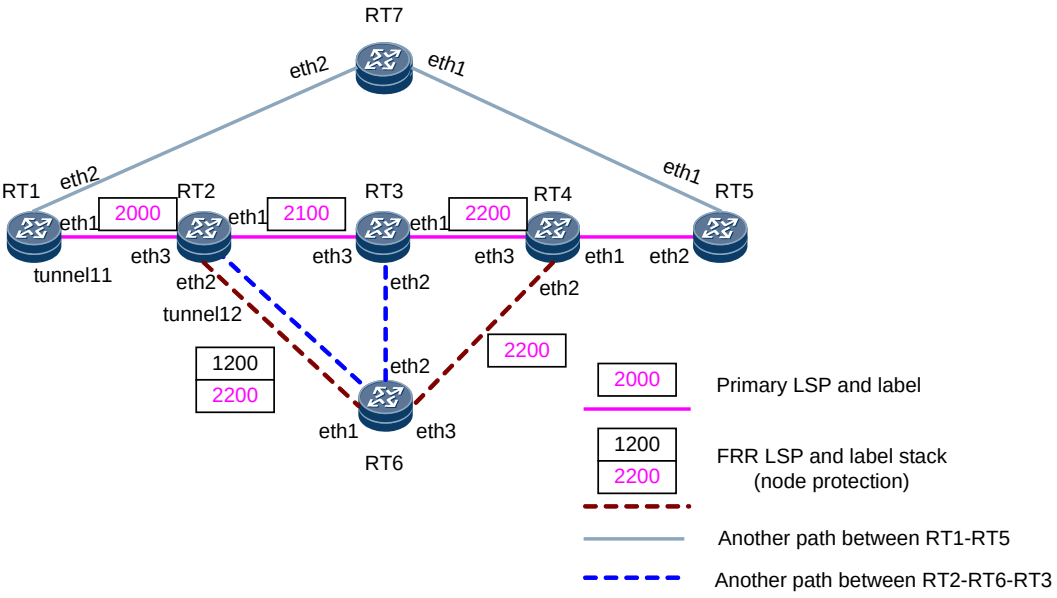


In Figure 10-8, the blue path is the primary LSP and the red path is the bypass LSP. When the link between RTB and RTC fails or when RTC fails, the data on the primary LSP is switched to the bypass LSP. The top layer of the packet header sent from RTB adopts the label assigned to RTB by RTF, and the egress label of RTC is also added to the label stack as the lower layer. The packet on the RTB-RTF-RTD LSP carries two labels. After receiving the packet, RTD finds the label assigned to RTF by RTD, and continues to use the label assigned to RTC by RTD for forwarding the packet.

Implementation Process of FRR in the Bypass Mode

Figure 10-9 illustrates the implementation process of FRR in the bypass mode.

Figure 10-9 FRR in the bypass mode



- Creating the primary LSP

  The primary LSP is created in the same way as an ordinary LSP is. The head node (RT1) sends the RSVP PATH message to downstream nodes one by one

(RT1-RT2-RT3-RT4-RT5), and the end node (RT5) sends the RESV message to upstream nodes one by one. When processing the RESV message, each node assigns the label and reserves the resources for creating the LSP.

In the protocol draft, some flag bits in the SESSION_ATTRIBUT and RECORD_ROUTE objects are extended for FRR. The difference between the creating processes of the protected LSP and ordinary LSP lies in the processing of these flag bits. The flag bits added to the SESSION_ATTRIBUT object in the PATH message indicate whether the LSP needs partial protection, whether the label is recorded, and whether the bandwidth is protected.

The flag bits added to the RECORD_ROUTE object in the RESV message indicate whether the LSP is protected, whether the switching is enabled, whether the bandwidth is protected, and whether the node is protected.

The creating of the primary LSP is triggered through the manual configuration of a tunnel on the head node (RT1). Before the primary LSP is created, if the FRR attribute is specified for the LSP by a command, the partial protection flag will be added to and the label flag and the SE style flag will be recorded in the SESSION_ATTRIBUTE object in the RSVP PATH message. If bandwidth is also specified for the LSP, the RSVP will also add the bandwidth protection flag. After receiving the PATH message, through the local protection flag, the downstream node can determine that the LSP requires the FRR protection.

For the LSP that requires the FRR protection (determined according to the flag in the PATH message received), each node records the egress, LSR ID, and label of the RESV message in the RRO when sending the RESV message to the upstream node. Such information is passed on to each upstream node.

When receiving the RESV message for the first time, according to the information recorded in the RRO, each node selects a proper bypass LSP for the LSP to be protected (primary LSP). The process of selecting a proper bypass LSP for the primary LSP is called binding.

After the node performs the FRR binding calculation on the primary LSP, the node indicates whether the primary LSP has been protected in the RECORD_ROUTE object in the RESV message sent to the upstream node. If the primary LSP has been protected, the egress (eth1 of RT2) of the protected LSP and the egress (eth3 of RT2) of the RESV message are recorded. If the primary LSP is not protected, the corresponding flag bit in the RRO is reset, and only the egress (eth3 of RT2) of the RESV message is recorded. Binding calculation is not performed on the egress. All the flag bits in the RRO sent from the egress to the upstream node are reset.

The primary LSP requiring the FRR protection is created in a similar way to an ordinary LSP. The differences are that, in the creating process of the primary LSP, the binding calculation is added, and related flag bits and sub-objects are added to the PATH and RESV messages.

- Creating the bypass LSP

  A bypass LSP can be created in two modes: the manual mode and the automatic mode.

  In the manual mode, after a tunnel without the FRR attribute is specified for protecting a physical interface, the LSP corresponding to this tunnel becomes a bypass LSP. A manual bypass LSP (tunnel12 on RT2) is configured manually on the PLR (RT2). The configuration of a manual bypass LSP is similar to an ordinary LSP. The difference is that the bypass LSP cannot be configured with the FRR attribute. In other words, a bypass LSP cannot be a primary LSP at the same time. An LSP cannot be protected by itself.

  The automatic mode of the bypass LSP simplifies the configuration of the manual mode. In the automatic mode, when the primary LSP requires the FRR protection, the PLR can select or automatically create a bypass LSP for protecting this primary LSP. A bypass

LSP can protect multiple primary LSPs in so far as it meets the requirements of these primary LSPs.

A bypass LSP can protect multiple physical interfaces, but it cannot protect its own egress.

FRR can implement link protection or node protection. In the configuration of the bypass LSP, the links or nodes to be protected should be planned, and whether the link protection mode or node protection mode is to be adopted should be determined. Generally, node protection is a superior mode because it can protect the protected nodes and the links between the PLR and the protected nodes. If conditions permit, the customer tends to require node protection. Huawei device provides flexible protection modes. When node protection fails, the protection mode supported by Huawei device can automatically shift to link protection. When node protection becomes valid again, node protection will be adopted.

The bandwidth of the bypass tunnel is generally used for protecting the primary LSP. All the resources of the bypass tunnel are used only after the switching occurs. Make sure that the configured bandwidth of the bypass LSP is equal to or greater than the sum of the bandwidth required by all the protected LSPs. Otherwise, after FRR takes effect, the bypass LSP will fail to provide the protection that meets the service quality requirements.

A bypass LSP is generally in the idle state and does not carry data. If the bypass tunnel is required to forward data as well as protecting the primary LSP, sufficient bandwidth should be configured.

- Binding calculation

  Binding can refer to specifying a bypass tunnel for protecting a physical interface. Then, the bypass tunnel can be said to be bound to the physical interface.

  Binding can also refer to selecting a proper bypass LSP for protecting a primary LSP. Then, the primary LSP can be said to be bound to the bypass LSP. The binding calculation is a process of binding a primary LSP to the bypass LSP. The result derived from the binding calculation is the necessary data to be forwarded in the switching, such as the interface of the bypass tunnel, the egress and NHLFE of the bypass LSP, and the label assigned by the MP. If the binding calculation is successful, the node sends the RESV message to inform the upstream node that the primary LSP has been protected.

  The binding calculation must be completed before the switching occurs. In the following conditions, binding calculation is triggered:

  - When a primary LSP is created
  - When the system periodically calculates the binding relations of all the LSPs whose egress is the protected physical interface

  The binding calculation always uses the known information of a primary LSP to traverse the bypass LSPs on the egress through which the primary LSP is protected, thus to find a most suitable bypass LSP. If automatic bypass LSP is supported, when a suitable bypass LSP is not found, the system will automatically try to create a bypass LSP for protecting the primary LSP.

  When the primary LSP is created, the interface address of each node is recorded. The CSPF can obtain the corresponding LSR ID according to the interface address. Hence, the LSR ID of the next hop (NHP) or next next hop (NNHOP) of the primary LSP is known. When the primary LSP is created, the RRO records the LSR ID of each hop. If the egress LSR ID and the NHP LSR ID of a bypass LSP are the same, link protection can be realized; if the egress LSR ID and NNHOP LSR ID of a bypass LSP are the same, node protection can be realized.

  If the bandwidth of a primary LSP is 0, it can be protected only by a bypass LSP whose bandwidth is 0. After a primary LSP comes into the protection of a bypass LSP whose

bandwidth is 0, the protection count of this bypass LSP is plus 1. If the bandwidth of a primary LSP is not 0, it can be protected only by a bypass LSP with sufficient remaining bandwidth. The initial remaining bandwidth of a bypass LSP whose bandwidth is not 0 is the configured value. Each time a primary LSP comes into the protection of the bypass LSP, the remaining bandwidth of the bypass LSP is minus the bandwidth of the primary LSP.

When multiple bypass LSPs are available for protecting a primary LSP, the following priority is adopted:

- Node protection is prior to link protection.
- If the bandwidth of the primary LSP is 0, a bypass LSP whose bandwidth is 0 is selected. If the bandwidth of the primary LSP is not 0, the bypass LSP whose remaining bandwidth is equal to or greater than the bandwidth of the primary LSP is selected.

The result derived from the binding calculation contains the following items, which are used for sending the data and signaling message from the bypass tunnel after the switching.

- Protection type (link protection or node protection), and the LSR ID of the MP.
- The label assigned to the last hop by the MP. This label is the label corresponding to the MP LSR ID in the RRO of the primary LSP.
- Egress and NHLFE of the bypass tunnel.

The binding calculation result is saved and can be immediately used when partial failure occurs. This is why MPLS TE FRR can respond quickly to failure.

- Failure detection

  The purpose of failure detection is to detect the failure of a link (RT2-RT3) or a node (RT3) as soon as possible so that switching can be triggered to reduce packet loss. Failure detection does not specifically distinguish between a link and a node, and the result of failure detection is presented as "interface failure" (eth1 of RT2).

  The "interface failure" triggers the FRR switching on all the LSPs that use the interface as the egress. If an LSP has been determined by the binding calculation to be in the link protection, the LSP will switch to link protection. If the actual failure is a node failure, the switching fails. As a result, the LSP is deleted. If an LSP has been determined by the binding calculation to be in the node protection, the LSP will switch to node protection. If the actual failure is a link failure and even if the next hop is available, the next hop will be skipped by the bypass tunnel.

  Certain link or node failures can be detected by the link layer protocol. The detection speed of the link layer protocol is directly related to the interface type. Other link or node failures are detected through the hello mechanism of the RESV. The detection speed of the hello mechanism is relatively slow.

  The hello function can be enabled on each physical interface that needs protection and on its interconnected interface. Then, the hello message and the response will be sent between the two routers periodically. In case of a link or node failure, the hello message or the response is lost. When the messages are lost for three successive times, it is regarded that a failure occurs.

- Switching

  Switching refers to adopting the bypass LSP for sending the data and RSVP messages that used to be sent through the primary LSP. When the interface (eth1 of RT2) is shut down by a command or when "interface failure" (of eth1 of RT2) is detected through the failure detection mechanism, switching is triggered. In the switching, the data and signaling of the protected LSPs on the faulty interface are switched to the bypass LSP for sending, and the upstream node is informed that switching occurs.

During the binding calculation by the forwarding component involved in the switching, the inner label (2200) required for the forwarding has been saved in the NHLFE. Now, it only needs to indicate that the LSP has been switched, and the data can be forwarded through the bypass tunnel.

Then, the node will respond to the switching event through the RESV message. For the LSP that has been bound to the bypass LSP, the node sends the upstream node the RSVP PathError message with the switching flag bit. The bypass tunnel is mainly used for temporary protection. The head node will properly process the LSPs that have been switched. If an LSP is not bound to the bypass LSP, the node directly sends the RSVP ResvTear message to inform the upstream node to delete the LSP.

- Maintenance of the LSP after the switching

  After the switching, the original link is not available. To prevent the LSP from being deleted after timeout, the information between the PLR (RT2) and the MP (RT4) needs to be refreshed through RSVP messages.

  After being modified, the PATH message is sent to the MP through the bypass tunnel (Tunnel12 of RT2). After receiving the PATH message, the MP confirms itself as an MP. Then, the RESV message is modified and forwarded to the PLR through the IP addresses of multicast hops (RT4-RT6-RT2).

  After the switching, the message sent from the PLR to the upstream node is also changed. That is, the address of the egress (eth2 of RT2) of the bypass LSP is added to the RRO.

  After the switching, the sending path of the PTEAR, RERR, RTEAR and PERR messages of the primary LSP are changed accordingly.

  After the switching in node protection, the protected node (RT3) may send the PATHTEAR message to the downstream node because the PATH message times out. In this case, the MP (RT4) ignores this message. In addition, in the switching, the MP sends the ResvTear message from the ingress (eth3 of RT4) of the original LSP. Thus, the protected node (RT3) will release the corresponding resource as soon as possible.

- Re-optimization

  Re-optimization refers to calculating the path for a created LSP at the preset intervals. According to the calculated path, the router initiates the creating of a new LSP. After the new LSP is created, the original LSP is deleted, and the data of the original LSP tunnel is switched to the new LSP for forwarding.

  Re-optimization can be configured for each LSP tunnel. After the LSP is created, re-optimization is enabled.

  In the case of FRR, another function of re-optimization is to restore the tunnel (Tunnel1 of RT1) protected by the bypass LSP to the normal state. This is because the FRR protection is temporary. Therefore, a tunnel with the FRR attribute is generally configured with re-optimization. When the primary LSP has not switched, a new LSP is created only when the path calculated through re-optimization is different from the original path. When the primary LSP has switched, a new LSP is created even when the path calculated through re-optimization is the same as the original path.

  A bypass LSP that has been bound to a physical interface can also be re-optimized. The bypass LSP, however, cannot be re-optimized if a primary LSP already switches to this bypass LSP. After a bypass LSP is re-optimized, the binding relations between the bypass LSP and the primary LSPs are refreshed.

  Before the primary LSP is switched, the data forwarding is the same as that of an ordinary LSP; after the primary LSP is switched to the bypass tunnel, the data is forwarded through the bypass tunnel to the MP.

  When the primary LSP is successfully bound to the bypass LSP, the NHLFE entry and the inner label (2200, the label assigned to the last upstream node by the MP) of the bypass LSP are recorded in the NHLFE entry of the primary LSP. In the switching, the

forwarding component sets the switching flag bit in the NHLFE entry of the primary LSP.

When the packet arrives at the PLR, the forwarding component searches for the NHLFE entry to the primary LSP. If switching has not occurred, the component performs label switching and data forwarding; if the switching flag bit is found in the NHLFE entry, the component continues searching for the NHLFE entry to the corresponding bypass LSP. After finding the NHLFE entry, the component adds inner label 2200 to the label stack, and performs forwarding according to the information of the NHLFE entry of the bypass LSP.

At the egress of the bypass tunnel (or at the last but one hop), inner label 2200 is removed from the label stack, and then MP can perform forwarding by using the original label 2200. The inner label may be used on different interfaces of the MP. Therefore, the MP must assign a label to each platform.

As previously mentioned, certain failures are detected at the link layer. After a failure is detected at the link layer, the forwarding component can reset the switching flag bit in the NHLFE entry of the primary LSP if the failure recovers before a corresponding failure occurs at the upper layer. Hence, the data of the primary LSP is still forwarded through the original path, and the switching flag in the RESV message is not processed.

One thing should be noted that, after the switching, the RSVP message from the PLR to the MP is sent through the bypass tunnel. In other words, the message is forwarded as a common IP packet through the MPLS tunnel. The RSVP message from the MP to the PLR is forwarded as a common IP packet.

10.6.2 TE Tunnel Protection Group

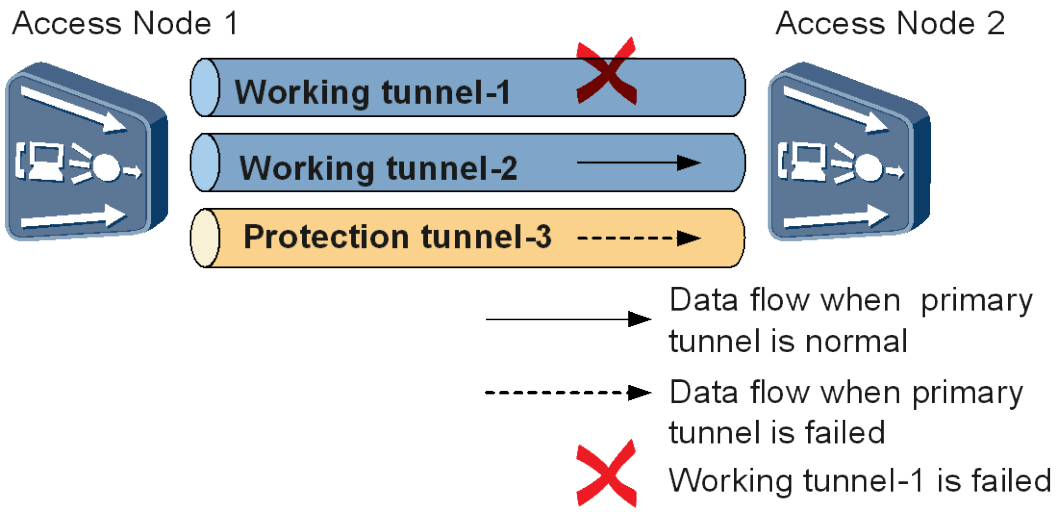
A tunnel protection group protects end-to-end MPLS TE tunnels. If a working tunnel in a protection group fails, traffic switches to a protection tunnel, minimizing traffic interruptions.

Related Concepts

As shown in the Figure 10-10, concepts related to a tunnel protection group are as follows:

- Working tunnel: a tunnel to be protected.
- Protection tunnel: a tunnel that protects a working tunnel.
- Protection switchover: switches traffic from a faulty working tunnel to a protection tunnel in a tunnel protection group, which improves network reliability.

Figure 10-10 Tunnel protection group



Primary tunnels tunnel-1 and tunnel-2, and the bypass tunnel tunnel-3 are established on the ingress Access Node shown in the Figure 10-10.

Tunnel-3 is specified as a protection tunnel for primary tunnels tunnel-1 and tunnel-2 on Access Node. If the configured fault detection mechanism on the ingress detects a fault in tunnel-1, traffic switches to tunnel-3. Access Node attempts to reestablish tunnel-1. If tunnel-1 is successfully established, traffic switches back to the primary tunnel.

Principle

Implementation

A TE tunnel protection group uses a configured protection tunnel to protect traffic on the working tunnel to improve tunnel reliability. To ensure the improved performance of the protection tunnel, the protection tunnel must exclude links and nodes through which the working tunnel passes during network planning.

Table 10-2 shows the implementation procedure of a tunnel protection group.

Table 10-2 Implementation procedure of a tunnel protection group

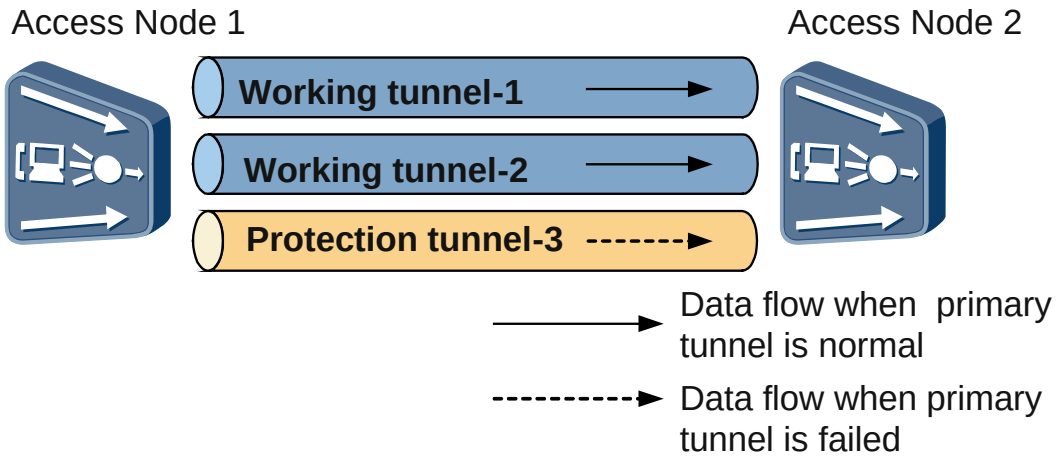
Sequence Number	Process	Description
1	Establishment	The working and protection tunnels must have the same ingress and destination address. The protection tunnel is established in the same procedure as a regular tunnel. The protection tunnel can use attributes that differ from those for the working tunnel. Ensure that the working and protection tunnels are established over different paths as much as possible.

Sequence Number	Process	Description
		NOTE • A protection tunnel cannot be protected or enabled with TE FRR. • Attributes for a protection tunnel can be configured independently of those for the working tunnel, which facilitates the network planning.
2	Binding between the working and protection tunnels	The protection tunnel is bound to the tunnel ID of the working tunnel so that the two tunnels form a tunnel protection group.
3	Fault detection	In addition to MPLS TE's own detection mechanism, MPLS OAM and BFD for CR-LSP are used to detect faults in a tunnel protection group to speed up protection switching.
4	Protection switching	The tunnel protection group supports either of the following protection switching modes: • Manual switching: Traffic is forcibly switched to the protection tunnel. • Automatic switching: Traffic automatically switches to the protection tunnel if the working tunnel fails. A time interval can be set for automatic switching.
5	Switchback	After a traffic switchover is implemented, the ingress attempts to reestablish the working tunnel. If the working tunnel is reestablished, the ingress can switch traffic back to the working tunnel or still forward traffic over the protection tunnel.

:Protection mode

A tunnel protection group works in either 1:1 or N:1 mode. The 1:1 mode enables a protection tunnel to protect only a single working tunnel. The N:1 mode enables a protection tunnel to protect more than one working tunnel.

Figure 10-11 N:1 protection mode



Differences Between CR-LSP Backup and a Tunnel Protection Group

CR-LSP backup and a tunnel protection group are both E2E protection mechanisms for MPLS TE. Table 10-3 shows the comparison between these two mechanisms.

Table 10-3 Comparison between CR-LSP backup and a tunnel protection group

Item	CR-LSP Backup	Tunnel Protection Group
Object to be protected	Primary and backup CR-LSPs are established on the same tunnel interface. A backup CR-LSP protects traffic on a primary CR-LSP.	One tunnel protects traffic over another tunnel in a tunnel protection group.
TE FRR	A primary CR-LSP supports TE FRR. A backup CR-LSP does not support TE FRR.	A working tunnel supports TE FRR. A protection tunnel does not support TE FRR.
LSP attributes	Primary and backup CR-LSPs have the same attributes, except for the TE FRR attribute.	The attributes of one tunnel in a tunnel protection group are independent of the attributes of the other tunnel. For example, a protection tunnel with no bandwidth can protect traffic on a working tunnel that has a bandwidth.
Protection mode	A 1:1 protection mode is supported. Each primary CR-LSP is protected by a backup CR-LSP.	An N:1 protection mode is supported. Many tunnels share one protection tunnel. If any protected tunnel fails, traffic switches to the protection tunnel.

10.6.3 CR-LSP Backup

CR-LSP backup techniques protect E2E MPLS TE tunnels. If the ingress detects that the primary CR-LSP is unavailable, the ingress switches traffic to a backup CR-LSP. After the primary CR-LSP recovers, traffic switches back.

Related Concepts

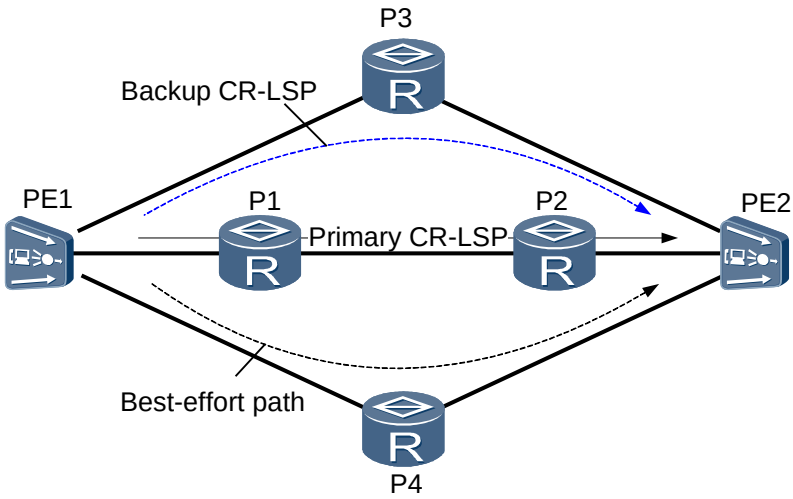
CR-LSP backup functions include hot standby, ordinary backup, and the best-effort path function. CR-LSP backup functions are as follows:

- Hot standby: A hot-standby CR-LSP is established immediately after a primary CR-LSP is created. If the primary CR-LSP fails, the hot-standby CR-LSP takes over traffic from the primary CR-LSP. After the primary CR-LSP recovers, traffic switches back.
- Ordinary backup: An ordinary backup CR-LSP can be established only after a primary CR-LSP fails. The ordinary backup CR-LSP takes over traffic if the primary CR-LSP fails. After the primary CR-LSP recovers, traffic switches back.
- Best-effort path

  If both the primary and backup CR-LSPs fail, a best-effort path is established and takes over traffic.

  For example, the primary CR-LSP is established over the path PE1 → P1 → P2 → PE2, and the backup CR-LSP is established over the path PE1 → P3 → PE2 shown in Figure 10-12. If both CR-LSPs fail, PE1 establishes a best-effort path PE1 → P4 → PE2 to take over traffic.

Figure 10-12 Best-effort path



NOTE

A best-effort path has no bandwidth reserved for traffic, but has an affinity and a hop limit configured as needed.

Principle

Implementation

The procedure of CR-LSP backup is as follows:

1. CR-LSP backup is deployed.

   Plan the paths, bandwidth values, and deployment modes. Table 10-4 lists CR-LSP backup deployment items.

Table 10-4 CR-LSP backup deployment

Item	Hot Standby	Ordinary Backup	Best-Effort Path
Path	Determine whether the primary and hot-standby CR-LSPs entirely or partially overlap. A hot-standby CR-LSP can be established over an explicit path. A hot-standby CR-LSP supports the following attributes: • Explicit path • Affinity • Hop limit	Allowed to use the path of the primary CR-LSP in all scenarios. An ordinary backup CR-LSP supports the following attributes: • Explicit path • Affinity • Hop limit	Automatically calculated by the ingress. A best-effort path supports the following attributes: • Affinity • Hop limit
Bandwidth	A hot-standby CR-LSP and a primary CR-LSP have the same bandwidth by default.	An ordinary backup CR-LSP and a primary CR-LSP have the same bandwidth.	A best-effort path is only a protection path that does not have reserved bandwidth.
Configuration combination	A hot-standby CR-LSP can be used together with a best-effort path.	An ordinary CR-LSP can only be used alone.	–

2. Fault detection is implemented.

   CR-LSP backup supports the RSVP-TE fault advertisement mechanism, who sends signaling packets to detect faults at a low speed.

3. A traffic switchover is implemented.

   If a primary CR-LSP fails, the ingress attempts to switch traffic from the primary CR-LSP to a hot-standby CR-LSP. If the hot-standby CR-LSP is unavailable, the ingress

attempts to switch traffic to an ordinary backup CR-LSP. If the ordinary backup CR-LSP is unavailable, the ingress attempts to switch traffic to a best-effort path.

4. A traffic switchback is implemented.

   Traffic switches back to a path based on the available CR-LSPs. Traffic will switch first to the primary CR-LSP, which has the highest priority. If the primary CR-LSP is unavailable, traffic will switch to the hot-standby CR-LSP. The ordinary CR-LSP has the lowest priority.

Overlapping Path for a Hot-standby CR-LSP

The overlapping path function can be configured for a hot-standby CR-LSP. The path of the hot-standby CR-LSP can overlap the path of a primary CR-LSP in all scenarios.

Coexistence of CR-LSP Backup and TE FRR

1. CR-LSP backup functions can be used together with TE FRR.
   - Hot standby and TE FRR: If TE FRR detects a link fault, traffic switches to a TE FRR bypass CR-LSP and then to a hot-standby CR-LSP.
   - Ordinary backup and TE FRR: If TE FRR detects a link fault, traffic switches to a TE FRR bypass CR-LSP. If both the primary and TE FRR bypass CR-LSPs fail, an ordinary backup CR-LSP is established and takes over traffic.
2. CR-LSP backup can be associated with TE FRR.

   The association improves tunnel security. The association provides the following functions based on backup modes:
   - Association between an ordinary backup CR-LSP and a TE FRR bypass CR-LSP provides the following functions:

     If a protected link or node fails, traffic switches to a bypass CR-LSP. The ingress attempts to reestablish the primary CR-LSP, while attempting to establish an ordinary backup CR-LSP.

     If the ordinary backup CR-LSP is established successfully before the primary CR-LSP is restored, traffic switches to the ordinary backup CR-LSP.

     After the primary CR-LSP recovers, traffic switches back to the primary CR-LSP.

     If the ordinary backup CR-LSP fails to be established, and the primary CR-LSP does not recover, the traffic still passes through the bypass CR-LSP.
   - Association between a hot-standby CR-LSP and a TE FRR bypass CR-LSP provides the following functions:

     If a hot-standby CR-LSP is Up and a protected link or node fails, traffic switches to a TE FRR bypass CR-LSP and then immediately switches to the hot-standby CR-LSP. At the same time, the ingress attempts to restore the primary CR-LSP.

     If the hot-standby CR-LSP is Down, the traffic switching procedure is the same as that when the ordinary backup is used.

   Association between ordinary backup CR-LSPs and TE FRR is recommended. An ordinary backup CR-LSP without additional bandwidth needed is established only after the primary CR-LSP enters the FRR-in-use state. Although the primary CR-LSP is Up, the system attempts to establish a hot-standby CR-LSP with additional bandwidth needed.

10.7 Configuring the MPLS Service

This topic describes the MPLS technology and how to configure the MPLS service on the MA5600T/MA5603T/MA5608T.

Basic concept

- The path that an FEC traverses in an MPLS network is called LSP. The LSP, whose function is the same as the virtual circuit in ATM and frame relay, is a unidirectional path from the ingress to the egress. Each node on the LSP is an LSR.
- The static LSP is the label forwarding path manually set up for label distribution to each FEC.
- The dynamic LSP is the label forwarding path dynamically established through the label distribution protocol (LDP or RSVP-TE).

Configuration logic

In the MPLS configuration, the core is to configure the LSP and the second is to configure fault detection and protection for the LSP. At the same time, According to the protocol for creating LSPs, LSPs are categorized as static LSP, LDP LSP, and RSVP-TE LSP.

Therefore, configure MPLS as follows:

1. Configure LSPs.
   - Configure a static LSP.
   - Configure an LDP LSP.
   - Configure an RSVP-TE LSP.
2. Configure LSP protection. Configure the MPLS OAM.

10.7.1 Configuring the Static LSP

Static LSP is configured manually. A static LSP can work in the normal state only when all the LSRs along the static LSP are configured.

Prerequisites

1. The IP address of the loopback interface must be configured.
2. The LSR ID must be configured.
3. The global MPLS, VLAN MPLS, and VLAN interface MPLS must be enabled.
4. A static or dynamic route must be successfully configured on each device in the network (so that LSRs can reach each other through the IP route).

Context

The administrator needs to manually distribute labels to each LSR when configuring the static lsp. Principle: The out label value of a node must be equal to the in label value of its next node. LSRs on a static LSP cannot perceive the entire LSP. Therefore, static LSP is a local concept.

The MA5600T/MA5603T/MA5608T can function as a label switching edge router (LER) or a label switching router (LSR). According to the position of the LER or LSR in a network, the

configuration of the static LSP involves the ingress configuration, transit node configuration, and egress configuration.

An LSP corresponds to a unidirectional forwarding path. To ensure bidirectional communication of the MPLS service, two static LSPs are required. The two LSPs have opposite directions. Their ingress and egress are reverse. Their transit nodes can be the same or different according to the networking requirements, or even free of being configured.

Procedure

- When the MA5600T/MA5603T/MA5608T functions as an LER, configure the static LSP as follows:
  a. Run the **static-lsp ingress** command to configure the ingress parameters of a static LSP.

     An LER is generally located at the edge of an MPLS network. The PE or PTN device can be considered an LER.

     Format:

     static-lsp ingress { *lsp-name* | **tunnel-interface** **tunnel** *tunnel-id* } **destination** *ip-addr* **nexthop** *ip-addr* **out-label** *out-label*

     - You can create a static LSP by using the LSP name or the tunnel. To create a static LSP by using the tunnel, you must run the **interface tunnel** command to create a tunnel interface and then configure its attributes.
     - **destination** *ip-addr*: Indicates the destination IP address of the LSP, that is, the loopback interface IP address of the PE or PTN device.
     - **nexthop** *ip-addr*: Indicates the next hop IP address, that is, the VLAN interface IP address of the adjacent LSR.
     - **out-label** *out-label*: Indicates the out label value, which must be the same as the in label value of the downstream LSR.

  b. Run the **static-lsp egress** command to configure the egress parameters of a static LSP.

     Format:

     static-lsp egress *lsp-name* **incoming-interface** **vlanif** *vlanid* **in-label** *in-label*[**lsrid** *ingress-lsr-id* **tunnel-id** *tunnel-id*]

     - In the egress configuration of a static LSP, only a VLAN interface can be used as the ingress interface.
     - **in-label** *in-label*: Indicates the in label value of the egress, which must be the same as the out label value of the upstream LSR.

  c. Run the **display mpls static-lsp** command to query the configuration of a static LSP.
- When the MA5600T/MA5603T/MA5608T functions as an LSR, configure the static LSP as follows:
  a. Run the **static-lsp transit** command to configure the transit node parameters of a static LSP.

     An LSR is generally located in the middle of an MPLS network. The P device can be considered an LSR that forwards MPLS labels.

     Format:

     static-lsp transit *lsp-name* **incoming-interface** *interface-type interface-number* **in-label** *in-label* **nexthop** *next-hop-address* **out-label** *out-label*

- The ingress interface of the transit node on a static LSP can only be the VLAN interface, that is, the VLAN interface of the upstream egress.
- **in-label** *in-label*: Indicates the in label value of the transit node, which must be the same as the out label value of the upstream ingress.
- **nexthop** *next-hop-address*: Indicates the next hop IP address, that is, the VLAN interface IP address of the adjacent LSR.
- **out-label** *out-label*: Indicates the out label value of the transit node, which must be the same as the in label value of the downstream LSR.

NOTICE

Because the LSP is unidirectional, you must configure the transit node parameters twice with opposite directions to ensure bidirectional communication of the MPLS service.

b. Run the **display mpls static-lsp** command to query the configuration of a static LSP.

----End

Example

When the MA5600T/MA5603T/MA5608T functions as an LER, to configure the ingress and egress of a static LSP, set the parameters as follows:

- Ingress node name of the static LSP: lsp1; egress name of the static LSP: lsp2
- IP address of local VLAN interface 100: 100.1.1.2/24
- Destination IP address of the LSP: 3.3.3.3/32
- Out label: 8200; in label: 8300
- Next hop IP address: 100.1.1.3

```
huawei(config)#static-lsp ingress lsp1 destination 3.3.3.3 32 nexthop 100.1.1.3
out-label 8200
huawei(config)#static-lsp egress lsp2 incoming-interface vlanif 100 in-label 8300
huawei(config)#display mpls static-lsp
{ <cr>|exclude<K>|include<K>|string<S><Length 1-19>|verbose<K> }:

  Command:
        display mpls static-lsp
TOTAL         :       2      STATIC LSP(S)
UP            :       0      STATIC LSP(S)
DOWN          :       2      STATIC LSP(S)
Name              FEC              I/O Label    I/O If              Status
lsp1              3.3.3.3/32         NULL/8200    -/-                Down
lsp2              -/-              8300/NULL    vlanif100/-         Down
```

When the MA5600T/MA5603T/MA5608T functions as an LSR, to configure the transit node parameters of a static LSP, set the parameters as follows:

- LSP name of the transit node in the positive direction: lsp1; LSP name of the transit node in the negative direction: lsp2

- IP address of local VLAN interface 100: 100.1.1.2/24
- IP address of local VLAN interface 200: 200.1.1.2/24
- Out label in the positive direction: 8200; in label in the positive direction: 8300
- Out label in the negative direction: 8200; in label in the negative direction: 8300
- Next hop IP address in the positive direction: 200.1.1.3
- Next hop IP address in the negative direction: 100.1.1.3

```
huawei(config)#static-lsp transit lsp1 incoming-interface vlanif 100 in-label 82
00 nexthop 200.1.1.3 out-label 8300
huawei(config)#static-lsp transit lsp2 incoming-interface vlanif 200 in-label 83
00 nexthop 100.1.1.2 out-label 8200
huawei(config)#display mpls static-lsp
{ <cr>|exclude<K>|include<K>|string<S><Length 1-19>|verbose<K> }:

  Command:
        display mpls static-lsp
TOTAL         :       2      STATIC LSP(S)
UP            :       0      STATIC LSP(S)
DOWN          :       2      STATIC LSP(S)
Name              FEC              I/O Label   I/O If              Status
lsp1              -/-              8200/8300   vlanif100/-          Down
lsp2              -/-              8300/8200   vlanif200/-          Down
```

10.7.2 Configuring the LDP LSP

Set up an MPLS LDP session between LSRs along the LSP. After the MPLS LDP session is set up, the LDP LSP is automatically created.

Prerequisites

1. The IP address of the loopback interface must be configured.
2. The LSR ID must be configured.
3. The VLAN for MPLS label forwarding must be created.
4. Global MPLS must be enabled.
5. A static or dynamic route must be successfully configured on each device in the network (so that LSRs can reach each other through the IP route).

Context

- The MA5600T/MA5603T/MA5608T supports LDP and RSVP-TE, both of which generate dynamic LSPs.
- LDP is a standard MPLS label distribution protocol defined by IETF. LDP, which is mainly used to distribute labels for the negotiation between LSRs to set up label switching paths (LSPs), regulates various types of information for the label distribution process, and the related processing. The LSRs form an LSP that crosses the entire MPLS domain according to the local forwarding table, which correlates the in label, network hop node, and out label of each specific FEC.

Procedure

Configure the MPLS LDP session.

The MPLS-LDP session is used for information exchange such as label mapping and release between LSRs. The MPLS-LDP session is classified into two types:

- Local LDP session: Two LSRs between which a session is set up are connected directly.
- Remote LDP session: Two LSRs between which a session is set up are not connected directly. Remote LDP sessions are mainly set up between nonadjacent LSRs. They can also be set up between adjacent LSRs.

NOTE

If local adjacency with the specified remote peer exists, remote adjacency cannot be set up; if remote adjacency exists and local adjacency is set up for the remote peer, the remote peer will be deleted. In other words, only one session can exist between two LSRs and a local LDP session takes priority over a remote LDP session.

- Configure the local LDP session.
  a. In the global config mode, run the **mpls ldp** command to enable global MPLS LDP.
  b. In the global config mode, run the **mpls vlan** command to enable the MPLS function of the VLAN.

NOTE

The VLAN 1 is the system default VLAN. All the upstream ports have been added to this VLAN by default. Do not use this VLAN as the MPLS VLAN or enable the MPLS function on this VLAN.

  c. Run the **interface vlanif** command to enter the VLAN interface mode.
  d. In the VLAN interface mode, run the **mpls** command to enable the MPLS function of the VLAN interface and run the **mpls ldp** command to enable the MPLS LDP function of the VLAN interface.
  e. Run the **quit** command to quit the VLAN interface mode.
- Configure the remote LDP session.
  a. In the global config mode, run the **mpls ldp** command to enable global MPLS LDP.
  b. Run the **mpls ldp remote-peer** command to create an LDP remote peer and then enter the remote peer mode.
  c. Run the **remote-ip** command to configure the IP address of the LDP remote peer.

NOTE

The IP address of the remote LDP peer should be the LSR ID of the remote LSR. When the LSR ID is used as the transmission address of a remote peer, two remote peers set up a TCP connection between them using the LSR ID as the transmission address.

  d. (Optional) Run the **mpls ldp advertisement** command to set the label distribution mode to DoD (downstream on demand) or DU (downstream unsolicited, default).

     In a network with a large scale, it is recommended to set the mode to DoD to reduce unnecessary MPLS forwarding entries.
  e. (Optional) Run the **remote-peer auto-dod-request** command to automatically use the DoD label distribution mode to request the label mapping information about the LSR IDs of all downstream remote peers.

     When the network has a large scale and many LDP remote peers, perform this configuration to maximally save system resources.

Step 1 (Optional) Configure the LDP MTU signaling function.

Run the **mtu-signalling** command to enable the sending of the MTU type, length, and value (TLV). This enables the LDP to automatically calculate and negotiate the minimum MTU value for all ports on each LSP. In this way, the MPLS determines the size of the MPLS forwarding packet at the ingress according to the minimum MTU, thereby avoiding the forwarding failure on transit nodes caused by oversize packets at the ingress.

By default, the LDP MTU signaling is enabled.

Step 2 (Optional) Configure the route trigger policy for setting up an LSP.

Run the **lsp-trigger host** command to configure the route trigger policy for setting up an LSP. The default route trigger policy is used to set up an LSP by triggering the LDP through the host address. To modify the default route trigger policy, run this command.

NOTE

It is recommended that you configure the route trigger policy for setting up an LSP to host (default), that is, the host route triggers the LDP to set up an LSP. In this way, the setup of useless LSPs can be prevented.

Step 3 (Optional) Configure the trigger policy set up by the transit LSP.

Run the **propagate mapping** command to filter certain routes received by the LDP by using the IP prefix table. Only the route that matches the specified IP prefix table is used by the local LDP for creating the transit LSP. By default, the LDP does not filter the received routes when creating the transit LSP.

Step 4 (Optional) Configure the LDP inter-domain extension function.

By default, LDP uses the full match mode to search for a route and set up an LSP; however, when the network scale is large and the LDP spans multiple IGP areas, the longest match mode must be used to search the routing table and set up an LSP accordingly. Run the **longest-match** command to configure the LDP inter-domain extension function.

Step 5 Query the relevant information about the LDP LSP configuration.

- Run the **display mpls ldp lsp** command to query the relevant information about the created LDP LSP.
- Run the **display mpls ldp session** command to check whether the created remote MPLS LDP session is in the normal (operational) state.
- Run the **display mpls interface** command to check whether the MPLS interface is in the normal (up) state.

----End

Example

To configure an LDP LSP between two adjacent LSRs by using VLAN interface 200 as the MPLS forwarding interface and using default values for other parameters, do as follows:

```
huawei(config)#mpls ldp
huawei(config-mpls-ldp)#quit
huawei(config)#mpls vlan 200
huawei(config)#interface vlanif 200
huawei(config-if-vlanif200)#mpls ldp
huawei(config-if-vlanif200)#quit
huawei(config)#display mpls interface vlanif 200
{ <cr>|verbose<K> }:

  Command:
        display mpls interface vlanif 200
Interface           Status    TE Attr   LSP Count  CRLSP Count Effective MTU
vlanif200           Down      Dis       0          0          1500
```

To configure an LDP LSP between two nonadjacent LSRs by configuring the local lsr-id to 3.3.3.3, configuring the remote lsr-id to 5.5.5.5, and using default values for other parameters, do as follows:

```
huawei(config)#mpls ldp
huawei(config-mpls-ldp)#quit
huawei(config)#mpls ldp remote-peer session1
huawei(config-mpls-ldp-remote-session1)#remote-ip 5.5.5.5
huawei(config-mpls-ldp-remote-session1)#quit
huawei(config)#display mpls ldp remote-peer
{ <cr>|peer-id<K>|string<S><Length 1-32>||<K> }:

 Command:
      display mpls ldp remote-peer

                    LDP Remote Entity Information
------------------------------------------------------------------------------
Remote Peer Name  : session1
Remote Peer IP    : 5.5.5.5          LDP ID       : 1.1.1.1:0
Transport Address : 1.1.1.1           Entity Status : Active

Configured Keepalive Hold Timer : 45 Sec
Configured Keepalive Send Timer : ---
Configured Hello Hold Timer     : 45 Sec
Negotiated Hello Hold Timer     : 45 Sec
Configured Hello Send Timer     : ---
Configured Delay Timer         : 10 Sec
Hello Packet sent/received      : 0/0
Label Advertisement Mode       : Downstream Unsolicited
Remote Peer Deletion Status     : No
Auto-config                  : ---
------------------------------------------------------------------------------
TOTAL: 1 Peer(s) Found.
```

10.7.3 Configure an RSVP-TE LSP

MPLS TE is a technology that integrates TE with MPLS. Through the MPLS TE technology, you can create an LSP tunnel to a specified path, to reserve resources and implement re-optimization.

Prerequisites

1. The IP address of the loopback interface must be configured.
2. The LSR ID must be configured.
3. The VLAN for MPLS label forwarding must be created.
4. Global MPLS and VLAN MPLS must be enabled.
5. The OSPF protocol must be successfully configured on each device in the network (the host route of each port must be successfully advertised).

Context

- To create constraint-based LSPs in MPLS TE, RSVP is extended. The extended RSVP signaling protocol is called the RSVP-TE signaling protocol.

- MPLS TE creates the LSP tunnel along a specified path through RSVP-TE and reserves resources. Thus, carriers can accurately control the path that traffic traverses to avoid the node where congestion occurs. This solves the problem that certain paths are overloaded and other paths are idle, utilizing the current bandwidth resources sufficiently. In addition, MPLS TE can reserve resources during the creation of LSP tunnels to ensure the QoS.

Procedure

Enable MPLS TE and RSVP-TE.

1. In the global config mode, run the **mpls** command to enter the MPLS mode.
2. In the MPLS mode, run the **mpls te** command to enable global MPLS TE, run the **mpls rsvp-te** command to enable global RSVP-TE, and run the **mpls te cspf** command to enable Constraint Shortest Path First (CSPF).
3. Run the **quit** command to quit the MPLS mode and run the **interface vlanif** command to enter the VLAN interface mode.
4. In the VLAN interface mode, run the **mpls** command to enable the VLAN interface MPLS, run the **mpls te** command to enable the VLAN interface MPLS TE, and run the **mpls rsvp-te** command to enable the VLAN interface RSVP-TE.

NOTE

- CSPF provides a way to select the path in an MPLS area. Enable CSPF before configuring other CSPF functions.
- It is recommended that you configure CSPF on all transit nodes lest the ingress cannot calculate the entire path.

Step 1 (Optional) Configure the line bandwidth.

To guarantee the bandwidth of the service transmitted on the MPLS TE tunnel, perform this operation.

1. In the VLAN interface mode, run the **mpls te bandwidth max-reservable-bandwidth** command to configure the maximum reservable bandwidth for the MPLS TE tunnel on the VLAN interface.
2. In the VLAN interface mode, run the **mpls te bandwidth** { **bc0** *bandwidth* | **bc1** *bandwidth* } command to configure the bandwidth that can be obtained from BC0 and BC1 of the VLAN interface when an MPLS TE tunnel is created.

NOTE

- BC0: Indicates the global pool bandwidth of an MPLS TE tunnel.
- BC1: Indicates the sub-pool bandwidth type of an MPLS TE tunnel. It is used to transmit services with higher priority and higher performance requirements.
- The bandwidth values must meet the following requirement: maximum reservable bandwidth ≥ BC0 bandwidth ≥ BC1 bandwidth.

Step 2 Enable MPLS TE for the OSPF area.

The MA5600T/MA5603T/MA5608T enables the MPLS TE to know the relevant dynamic TE attributes of each link by extending the OSPF protocol. The extended OSPF enables the link status entry to add TE attributes, such as link bandwidth and affinity attribute. Each router in the network collects all the TE information in OSPF area and generates traffic engineering database (TEDB).

1. In the global config mode, run the **ospf** command to start the OSPF process and enter the OSPF mode.
2. Run the **opaque-capability enable** command to enable the OSPF opaque capability.

After the opaque capability of the MA5600T/MA5603T/MA5608T is enabled, it can export TEDB information to neighbor devices.

3. Run the **area** command to enter the OSPF area mode and run the **mpls-te enable** command to enable the OSPF area TE.

Step 3 (Optional) Configure an MPLS TE explicit path.

An explicit path consists of a series of nodes, which constitute a vector path according to the configured sequence. The IP address in an explicit path is the IP address of the interface on the node. Generally, the loopback interface IP address on the egress is used as the destination IP address of the explicit path.

To specify a known path for a special traffic stream in the MPLS network, you can run the **explicit-path** command in the global config mode to configure an explicit path, and then run the **mpls te path explicit-path** command in the tunnel mode to specify the explicit path for the tunnel.

After an explicit path is created, you can run the **next hop**, **modify hop**, and **delete hop** command to add a next hop node, modify a node, and delete a node respectively for the explicit path.

Step 4 Configure an MPLS TE tunnel interface.

1. In global config mode, run the **interface tunnel** command to create a tunnel interface and enter the tunnel interface mode.
2. Run the **tunnel-protocol mpls te** command to configure the tunnel protocol to MPLS TE.
3. Run the **destination** *ip-address* command to configure the destination IP address of the tunnel. Generally, the egress LSR ID is used.
4. Run the **mpls te tunnel-id** command to configure the tunnel ID.
5. Run the **mpls te signal-protocol rsvp-te** command to configure the signaling protocol of the tunnel to RSVP-TE.
6. (Optional) Run the **mpls te bandwidth** command to configure the bandwidth for the tunnel. After the configuration is completed, only the VLAN interface that meets this bandwidth value can be selected as the node traversed by the MPLS TE tunnel path when the MPLS TE tunnel is created.

   If the MPLS TE tunnel is only used to change the data transmission path, you may not configure the tunnel bandwidth.
7. (Optional) Run the **mpls te path explicit-path** command to configure the explicit path used by the MPLS TE tunnel.

   If only the bandwidth used by the MPLS TE tunnel is limited but the transmission path is not limited, you may not configure the explicit path used by the MPLS TE tunnel.
8. Run the **mpls te commit** command to commit the current configuration of the tunnel.

Step 5 Check the configuration.

1. Run the **display mpls te cspf tedb** command to query the CSPF TEDB information.
2. Run the **display mpls te link-administration admission-control** command to check the CR LSP information allowed on the link, including the bandwidth and priority.
3. Run the **display mpls te tunnel** command to query details about a specified tunnel.
4. Run the **display mpls te tunnel path** command to query the path information about a tunnel on a local node.

5. Run the **display mpls te tunnel-interface** command to query the tunnel interface information about a local node.

----End

Example

To configure the RSVP-TE LSP from the MA5600T/MA5603T/MA5608T to the PTN, set the parameters as follows.

- Set the parameters on the MA5600T/MA5603T/MA5608T.
  - LSR-ID: 3.3.3.3
  - Layer 3 interface IP address of VLAN 20 for MPLS forwarding: 10.1.1.3/24
  - Maximum reservable bandwidth of the VLAN interface: 20480 kbit/s; BC0 bandwidth: 10240 kbit/s
  - OSPF process ID: 100; OSPF area ID: 1
  - MPLS TE tunnel ID: 10; tunnel interface ID: 10
  - Required BC0 bandwidth when an MPLS TE tunnel is created: 5120 kbit/s
  - Other parameters: default settings
- Set the LSR ID of the PTN to 5.5.5.5.

```
huawei(config)#interface loopback 0
huawei(config-if-loopback0)#ip address 3.3.3.3 32
huawei(config-if-loopback0)#quit
huawei(config)#mpls lsr-id 3.3.3.3
huawei(config)#mpls
huawei(config-mpls)#mpls te
huawei(config-mpls)#mpls rsvp-te
  //Configure the MPLS TE to use CSPF to calculate the shortest path to a node.
huawei(config-mpls)#mpls te cspf
huawei(config-mpls)#quit
huawei(config)#mpls vlan 20
huawei(config)#interface vlanif 20
  //Configure the IP address of the VLAN Layer 3 interface.
huawei(config-if-vlanif20)#ip address 10.1.1.3 24
  //Enable MPLS for the VLAN interface.
huawei(config-if-vlanif20)#mpls
  //Enable MPLS TE for the VLAN interface.
huawei(config-if-vlanif20)#mpls te
  //Enable MPLS RSVP-TE for the VLAN interface.
huawei(config-if-vlanif20)#mpls rsvp-te
huawei(config-if-vlanif20)#quit
huawei(config)#ospf 100
  //Enable the opaque capability to send the engineering data base information
    to peripheral devices.
huawei(config-ospf-100)#opaque-capability enable
huawei(config-ospf-100)#area 1
  //Enable MPLS TE for the OSPF area.
huawei(config-ospf-100-area-0.0.0.1)#mpls-te enable standard-complying
huawei(config-ospf-100-area-0.0.0.1)#quit
huawei(config-ospf-100)#quit
huawei(config)#interface vlanif 20
  //Configure the maximum reservable bandwidth of the Layer 3 interface.
```

```
huawei(config-if-vlanif20)#mpls te bandwidth max-reservable-bandwidth 20480
  //Configure the obtainable maximum bandwidth of the Layer 3 interface from BC0
    when the MPLS TE tunnel is created.
huawei(config-if-vlanif20)#mpls te bandwidth bc0 10240
huawei(config-if-vlanif20)#quit
huawei(config)#interface tunnel 10
  //Configure the link layer encapsulation protocol to MPLS TE for the tunnel interface,
    that is, configure the tunnel interface to work in the CR-LSP tunnel mode.
huawei(config-if-tunnel10)#tunnel-protocol mpls te
  //Configure the destination IP address of the MPLS TE tunnel.
huawei(config-if-tunnel10)#destination 3.3.3.3
  //Configure the MPLS TE tunnel ID, which, along with the LSR-ID,
    uniquely indicates an MPLS TE tunnel.
huawei(config-if-tunnel10)#mpls te tunnel-id 10
  //Configure the protocol of the MPLS TE tunnel to RSVP-TE.
huawei(config-if-tunnel10)#mpls te signal-protocol rsvp-te
  //Configure the global pool bandwidth required by the MPLS TE tunnel.
huawei(config-if-tunnel10)#mpls te bandwidth ct0 5120
  //Allow the MPLS TE tunnel to be bound to a VPN instance, that is, the MPLS TE tunnel
    can function as the outer tunnel of the PWE3 service.
huawei(config-if-tunnel10)#mpls te reserved-for-binding
huawei(config-if-tunnel10)#mpls te commit
huawei(config-if-tunnel10)#quit
```

10.7.4 Configuring the MPLS RSVP-TE FRR

The RSVP TE FRR technology is applied to the MPLS TE network for implementing partial network protection. Specifically, when a certain link or node in the network fails, the LSP configured with FRR can automatically switch the data to the protect link.

Prerequisites

1. The IP address of the loopback interface must be configured.
2. The LSR ID must be configured.
3. The VLAN for MPLS label forwarding must be created.
4. Global MPLS and VLAN MPLS must be enabled.
5. The OSPF protocol must be successfully configured on each device in the network (the host route of each port must be successfully advertised).

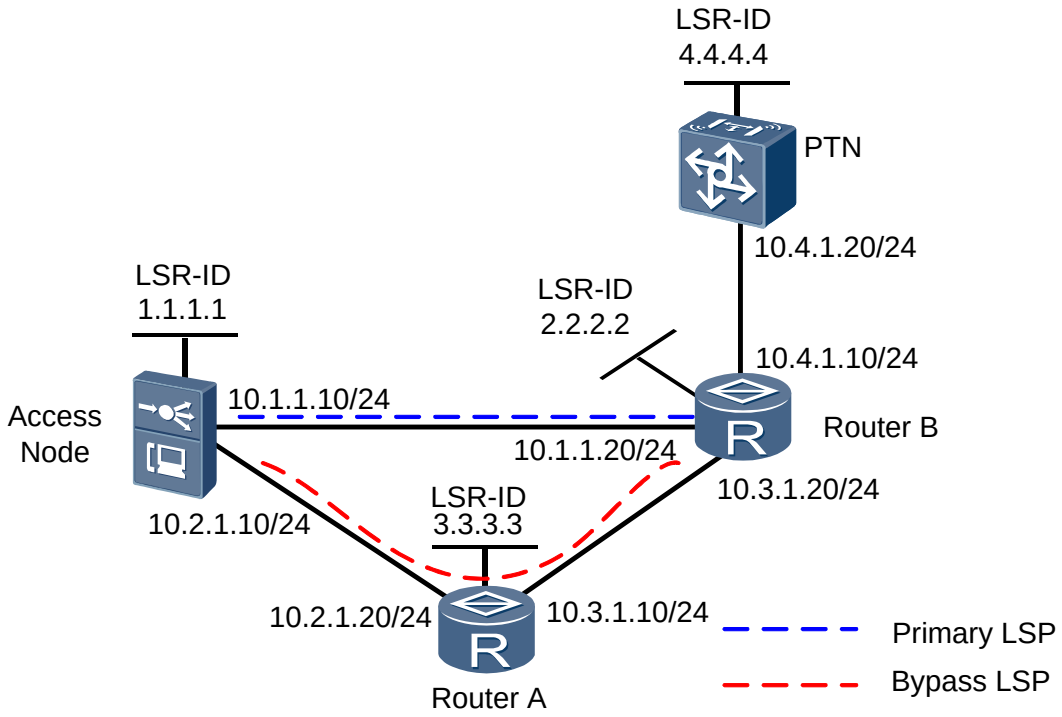
Context

- The implementation of the FRR is based on the extended RSVP-TE signaling. For the FRR, a protect tunnel is created in advance to protect the working tunnel. This prevents the broadcast delay of the notification between NEs and the duration for re-selecting the tunnel if the working tunnel fails. Therefore, the FRR can implement the second-level protection switchover.
- The MA5600T/MA5603T/MA5608T adopts the bypass mode (that is, using a protect path to protect multiple LSPs; the protect path is called the bypass LSP). Figure 10-13 shows the FRR function implemented in the bypass mode.

As shown in the figure, the blue dotted line indicates the primary LSP and the red dotted line indicates the by pass LSP. When the link or node between the MA5600T/MA5603T/MA5608T and Router B is faulty, services are switched to the bypass

link MA5600T/MA5603T/MA5608T->Router A->Router B. In this manner, the LSP is not affected by the link or node fault.

Figure 10-13 Example network of the MPLS RSVP-TE FRR



Procedure

Enable MPLS TE and RSVP-TE.

1. In the global config mode, run the **mpls** command to enter the MPLS mode.
2. In the MPLS mode, run the **mpls te** command to enable global MPLS TE, run the **mpls rsvp-te** command to enable global RSVP-TE, and run the **mpls te cspf** command to enable Constraint Shortest Path First (CSPF).
3. Run the **quit** command to quit the MPLS mode and run the **interface vlanif** command to enter the VLAN interface mode.
4. In the VLAN interface mode, run the **mpls** command to enable the VLAN interface MPLS, run the **mpls te** command to enable the VLAN interface MPLS TE, and run the **mpls rsvp-te** command to enable the VLAN interface RSVP-TE.

NOTE

- CSPF provides a way to select the path in an MPLS area. Enable CSPF before configuring other CSPF functions.
- It is recommended that you configure CSPF on all transit nodes lest the ingress cannot calculate the entire path.

Step 1 (Optional) Configure the line bandwidth.

To guarantee the bandwidth of the service transmitted on the MPLS TE tunnel, perform this operation.

1. In the VLAN interface mode, run the **mpls te bandwidth max-reservable-bandwidth** command to configure the maximum reservable bandwidth for the MPLS TE tunnel on the VLAN interface.
2. In the VLAN interface mode, run the **mpls te bandwidth** { **bc0** *bandwidth* | **bc1** *bandwidth* } command to configure the bandwidth that can be obtained from BC0 and BC1 of the VLAN interface when an MPLS TE tunnel is created.

NOTE

- BC0: Indicates the global pool bandwidth of an MPLS TE tunnel.
- BC1: Indicates the sub-pool bandwidth type of an MPLS TE tunnel. It is used to transmit services with higher priority and higher performance requirements.
- The bandwidth values must meet the following requirement: maximum reservable bandwidth ≥ BC0 bandwidth ≥ BC1 bandwidth.

Step 2 Enable MPLS TE for the OSPF area.

The MA5600T/MA5603T/MA5608T enables the MPLS TE to know the relevant dynamic TE attributes of each link by extending the OSPF protocol. The extended OSPF enables the link status entry to add TE attributes, such as link bandwidth and affinity attribute. Each router in the network collects all the TE information in OSPF area and generates traffic engineering database (TEDB).

1. In the global config mode, run the **ospf** command to start the OSPF process and enter the OSPF mode.
2. Run the **opaque-capability enable** command to enable the OSPF opaque capability.

   After the opaque capability of the MA5600T/MA5603T/MA5608T is enabled, it can export TEDB information to neighbor devices.
3. Run the **area** command to enter the OSPF area mode and run the **mpls-te enable** command to enable the OSPF area TE.

Step 3 Set up the primary tunnel on the MA5600T/MA5603T/MA5608T.

1. Configure the explicit path of the primary LSP.

   An explicit path consists of a series of nodes, which constitute a vector path according to the configured sequence. The IP address in an explicit path is the IP address of the interface on the node. Generally, the loopback interface IP address on the egress is used as the destination IP address of the explicit path.

   To specify a known path for a special traffic stream in the MPLS network, you can run the **explicit-path** command in the global config mode to configure an explicit path, and then run the **mpls te path explicit-path** command in the tunnel mode to specify the explicit path for the tunnel.

   After an explicit path is created, you can run the **next hop**, **modify hop**, and **delete hop** command to add a next hop node, modify a node, and delete a node respectively for the explicit path.
2. Configure the MPLS TE tunnel of the primary LSP.
   a. In global config mode, run the **interface tunnel** command to create a tunnel interface and enter the tunnel interface mode.
   b. Run the **tunnel-protocol mpls te** command to configure the tunnel protocol to MPLS TE.
   c. Run the **destination** *ip-address* command to configure the destination IP address of the tunnel. Generally, the egress LSR ID is used.
   d. Run the **mpls te tunnel-id** command to configure the tunnel ID.

e. Run the **mpls te signal-protocol rsvp-te** command to configure the signaling protocol of the tunnel to RSVP-TE.

f. (Optional) Run the **mpls te bandwidth** command to configure the bandwidth for the tunnel. After the configuration is completed, only the VLAN interface that meets this bandwidth value can be selected as the node traversed by the MPLS TE tunnel path when the MPLS TE tunnel is created.

If the MPLS TE tunnel is only used to change the data transmission path, you may not configure the tunnel bandwidth.

g. Run the **mpls te path explicit-path** command to configure the explicit path used by the MPLS TE tunnel.

h. Run the **mpls te commit** command to commit the current configuration of the tunnel.

3. Enable the FRR function of the tunnel.

Run the **mpls te fast-reroute** [**bandwidth**] command to enable TE FRR of the tunnel interface and allow bandwidth protection. By default, the FRR function is prohibited.

NOTE

Bandwidth protection configured through this command is used only for selecting the bypass tunnel policy. When the primary tunnel is faulty and needs to switch to a bypass tunnel, the bypass tunnel that meets the bandwidth requirement is preferred. If no bypass tunnel meets the bandwidth requirement, the primary tunnel selects an optimal bypass tunnel from the existing bypass tunnels.

Step 4 Set up a bypass LSP tunnel on the MA5600T/MA5603T/MA5608T.

1. Configure the explicit path of the bypass LSP.

a. In the global config mode, run the **explicit-path** command to configure the explicit path. In the tunnel mode, run the **mpls te path explicit-path** command to specify the explicit path for the tunnel.

b. Run the **next hop**, **modify hop**, and **delete hop** command to add a next hop node, modify a node, and delete a node respectively for the explicit path.

2. Configure the MPLS TE tunnel of the bypass LSP.

NOTICE

MPLS TE tunnel IDs of the primary and bypass LSPs cannot be the same.

a. In global config mode, run the **interface tunnel** command to create a tunnel interface and enter the tunnel interface mode.

b. Run the **tunnel-protocol mpls te** command to configure the tunnel protocol to MPLS TE.

c. Run the **destination** *ip-address* command to configure the destination IP address of the tunnel. Generally, the egress LSR ID is used.

d. Run the **mpls te tunnel-id** command to configure the tunnel ID.

e. Run the **mpls te signal-protocol rsvp-te** command to configure the signaling protocol of the tunnel to RSVP-TE.

f. (Optional) Run the **mpls te bandwidth** command to configure the bandwidth for the tunnel. After the configuration is completed, only the VLAN interface that meets this bandwidth value can be selected as the node traversed by the MPLS TE tunnel path when the MPLS TE tunnel is created.

If the MPLS TE tunnel is only used to change the data transmission path, you may not configure the tunnel bandwidth.

g. Run the **mpls te path explicit-path** command to configure the explicit path used by the MPLS TE tunnel.

h. Run the **mpls te commit** command to commit the current configuration of the tunnel.

3. Bind the bypass LSP tunnel to the protected interface.

a. In the tunnel mode, run the **mpls te bypass-tunnel** command to configure a bypass tunnel of the FRR.

NOTE

The total bandwidth of all LSPs that use bypass tunnels does not exceed the bandwidth of the primary tunnel. If multiple bypass tunnels exist, the system uses the best-fit algorithm to determine which bypass to use.

b. Run the **mpls te protected-interface** command to specify the interface to be protected by the bypass tunnel. When the interface is faulty, a bypass tunnel switching is triggered.

NOTE

One bypass tunnel can protect up to three interfaces, and MPLS TE must be enabled for the protected interfaces.

----End

Result

Enter the VLAN interface mode, and run the **shutdown** command to shut down the VLAN interface to disable the protected egress on the primary LSP. Then run the **display interface tunnel** command to query the status of the primary LSP on the MA5600T/MA5603T/MA5608T. You can see that the tunnel interface is still in the up state. Finally, run the **tracert lsp te tunnel** command to check the path traversed by the tunnel. You can see that the link is switched to the bypass tunnel.

Example

As shown in Figure 10-13, when the link or node between the MA5600T/MA5603T/MA5608T and Router B is faulty, services are switched to the standby link MA5600T/MA5603T/MA5608T->Router A->Router B. In this manner, the LSP is not affected by the fault of link or node. Set the parameters as follows:

- Set the parameters on the MA5600T/MA5603T/MA5608T.
  - LSR ID: 1.1.1.1
  - IP address of VLAN interface 10 connected to Router B: 10.1.1.10/24
  - IP address of VLAN interface 20 connected to Router A: 10.2.1.10/24
- Set the parameters on the Router B.
  - LSR ID: 2.2.2.2
  - IP address of the interface connected to the MA5600T/MA5603T/MA5608T: 10.1.1.20/24
  - IP address of the interface connected to Router A: 10.3.1.20/24
  - IP address of the interface connected to the PTN: 10.4.1.10/24
- Set the parameters on the Router A.
  - LSR ID: 3.3.3.3

  - IP address of the interface connected to the MA5600T/MA5603T/MA5608T: 10.2.1.20/24
  - IP address of the interface connected to Router B: 10.3.1.10/24
- Set the parameters on the PTN.
  - LSR ID: 4.4.4.4
  - IP address of the interface connected to Router B: 10.4.1.20/24

```
  //Configure the LSR-ID.
huawei(config)#interface loopback 0
huawei(config-if-loopback0)#ip address 1.1.1.1 32
huawei(config-if-loopback0)#quit
huawei(config)#mpls lsr-id 1.1.1.1
  //Enable RSVP-TE.
huawei(config)#mpls
huawei(config-mpls)#mpls te
huawei(config-mpls)#mpls rsvp-te
huawei(config-mpls)#mpls te cspf
huawei(config-mpls)#quit
  //Configure the IP address of VLAN interface 10 and enable RSVP-TE of the VLAN interface.
huawei(config)#vlan 10 standard
huawei(config)#mpls vlan 10
huawei(config)#interface vlanif 10
huawei(config-if-vlanif10)#ip address 10.1.1.10 24
huawei(config-if-vlanif10)#mpls
huawei(config-if-vlanif10)#mpls te
huawei(config-if-vlanif10)#mpls rsvp-te
huawei(config-if-vlanif10)#quit
  //Configure the IP address of VLAN interface 20 and enable RSVP-TE of the VLAN interface.
huawei(config)#vlan 20 standard
huawei(config)#mpls vlan 20
huawei(config)#interface vlanif 20
huawei(config-if-vlanif20)#ip address 10.2.1.10 24
huawei(config-if-vlanif20)#mpls
huawei(config-if-vlanif20)#mpls te
huawei(config-if-vlanif20)#mpls rsvp-te
huawei(config-if-vlanif20)#quit
  //Configure OSPF TE.
huawei(config)#ospf 100
huawei(config-ospf-100)#opaque-capability enable
huawei(config-ospf-100)#area 0
huawei(config-ospf-100-area-0.0.0.0)#mpls-te enable standard-complying
huawei(config-ospf-100-area-0.0.0.0)#quit
huawei(config-ospf-100)#quit
  //Configure the explicit path of the primary LSP.
huawei(config)#explicit-path pri-path
huawei(config-explicit-path-pri-path)#next hop 10.1.1.20
huawei(config-explicit-path-pri-path)#next hop 10.4.1.20
huawei(config-explicit-path-pri-path)#quit
  //Configure the MPLS TE tunnel of the primary LSP.
huawei(config)#interface tunnel 10
huawei(config-if-tunnel10)#tunnel-protocol mpls te
huawei(config-if-tunnel10)#destination 2.2.2.2
huawei(config-if-tunnel10)#mpls te tunnel-id 10
huawei(config-if-tunnel10)#mpls te signal-protocol rsvp-te
huawei(config-if-tunnel10)#mpls te path explicit-path pri-path
```

```
huawei(config-if-tunnel10)#mpls te fast-reroute
huawei(config-if-tunnel10)#mpls te commit
huawei(config-if-tunnel10)#quit
  //Configure the explicit path of the bypass LSP.
huawei(config)#explicit-path bypass-path
huawei(config-explicit-path-bypass-path)#next hop 10.2.1.20
huawei(config-explicit-path-bypass-path)#next hop 10.3.1.20
huawei(config-explicit-path-bypass-path)#quit
  //Configure the MPLS TE tunnel of the bypass LSP.
huawei(config)#interface tunnel 20
huawei(config-if-tunnel20)#tunnel-protocol mpls te
huawei(config-if-tunnel20)#destination 2.2.2.2
huawei(config-if-tunnel20)#mpls te tunnel-id 20
huawei(config-if-tunnel20)#mpls te signal-protocol rsvp-te
huawei(config-if-tunnel20)#mpls te path explicit-path bypass-path
huawei(config-if-tunnel20)#mpls te bypass-tunnel
huawei(config-if-tunnel20)#mpls te protected-interface vlanif 10
huawei(config-if-tunnel20)#mpls te commit
huawei(config-if-tunnel20)#quit
```

10.7.5 Configuring the MPLS OAM

The MPLS OAM function uses an effective OAM mechanism to detect whether an LSP is normal and report an alarm in time when an LSP fault occurs. In addition, the MPLS OAM function features a complete protection switching mechanism, which triggers a switchover when a defect at the MPLS layer is detected to minimize the data loss.

Context

Through the MPLS OAM mechanism, the MA5600T/MA5603T/MA5608T can effectively detect, confirm, and locate internal defects at the MPLS layer of a network. Then, the system reports and handles the defects. In addition, the system provides a mechanism for triggering 1:1 protection switching when a fault occurs.

The basic process of the MPLS OAM connectivity check and protection switching is as follows:

1. The source transmits the CV/FFD packets to the destination through the detected LSP.
2. The destination checks the correctness of the type and frequency carried in the received detection packets and measures the number of correct and errored packets that are received within the detection period to monitor the connectivity of the LSP in real time.
3. After detecting a defect, the destination transmits the BDI packets that carry the defect information to the source through the backward path.
4. The source learns about the status of the defect, and triggers the corresponding protection switching when the protect group is correctly configured.

Configure the MPLS OAM as follows:

1. Configure the active LSP at the source end (ingress).
2. Configure the standby LSP at the source end.
3. Create a tunnel protect group.
4. Enable the MPLS OAM function at the source end.
5. Configure the backward LSP at the destination end (egress).
6. Enable the MPLS OAM function at the destination end.

NOTE

If only the MPLS OAM connectivity check needs to be enabled and 1:1 protection is not required for the LSP, you need not configure the standby LSP or the tunnel protect group at the source end.

Configuration Example for Detection of MPLS OAM for Static LSP Connectivity

This topic describes how to configure the function of MPLS OAM to detect the static LSP connectivity.

Prerequisites

Before the configuration, make sure that:

- Set the IP addresses and the masks of the ports based on the example network. After that, LSRs can ping the peer LSRs.
- A static or dynamic route must be successfully configured on each device in the network (so that LSRs can reach each other through the IP route).

Networking

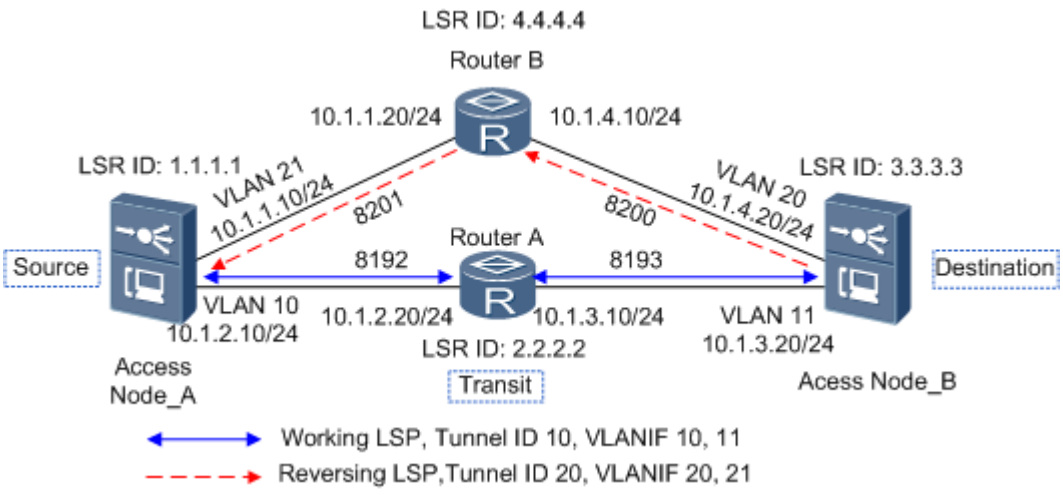
Figure 10-14 shows an example network of configuring MPLS OAM to detect the static LSP connectivity.

1. Source end MA5600T/MA5603T/MA5608T_A sends CV/FFD detection packets to the destination end through the detected LSP (MA5600T/MA5603T/MA5608T_A->Router A->MA5600T/MA5603T/MA5608T_B).
2. After detecting a defect, the destination transmits the BDI packets that carry the defect information to the source through the backward LSP (MA5600T/MA5603T/MA5608T_B->Router B->MA5600T/MA5603T/MA5608T_A). This enables the source end to obtain the defect status in time.

NOTE

To facilitate description of the MPLS OAM application, the MA5600T/MA5603T/MA5608T is used at both the source end and destination end as an example. In the actual application, the MA5600T/MA5603T/MA5608T at one end may be replaced by a device that supports MPLS OAM such as a PTN device, but their implementation principles are the same.

Figure 10-14 Example network of detection of MPLS OAM for static LSP connectivity



Data Plan

Table 10-5 provides the data plan for detection of MPLS OAM for static LSP connectivity.

Table 10-5 Data plan for detection of MPLS OAM for static LSP connectivity

Item	Data
MA5600T/MA5603T/MA5608T_A	LSR ID: 1.1.1.1
	Port: 0/19/0 IP address of VLAN interface 10 connected to Router A: 10.1.2.10/24 Tunnel ID: 10; tunnel interface ID: 10 Out label value of the LSP ingress: 8192 In label value of the LSP egress: 8193
	Port: 0/19/1 IP address of VLAN interface 21 connected to Router B: 10.1.1.10/24
	Static LSP: Router A to MA5600T/MA5603T/MA5608T_B
MA5600T/MA5603T/MA5608T_B	LSR ID: 3.3.3.3
	Port: 0/19/0 IP address of VLAN interface 11 connected to Router A: 10.1.3.20/24
	Port: 0/19/1 IP address of VLAN interface 20 connected to Router B: 10.1.4.20/24 Tunnel ID: 20; tunnel interface ID: 20 Out label value of the LSP ingress: 8200 In label value of the LSP egress: 8201
	Static LSP: Router B to MA5600T/MA5603T/MA5608T_A
Router A	LSR ID: 2.2.2.2
	IP address of the interface connected to the MA5600T/MA5603T/MA5608T_A: 10.1.2.20/24
	IP address of the interface connected to the MA5600T/MA5603T/MA5608T_B: 10.1.3.10/24
Router B	LSR ID: 4.4.4.4
	IP address of the interface connected to the MA5600T/MA5603T/MA5608T_A: 10.1.1.20/24
	IP address of the interface connected to the MA5600T/MA5603T/MA5608T_B: 10.1.4.10/24

Procedure

- **Configure source end MA5600T/MA5603T/MA5608T_A.**
  a. Configure the loopback interface.

```
huawei(config)#interface loopback 0
huawei(config-if-loopback0)#ip address 1.1.1.1 32
huawei(config-if-loopback0)#quit
```

  b. Enable the basic MPLS and MPLS TE.
     i. Enable the basic MPLS and MPLS TE globally.

```
huawei(config)#mpls lsr-id 1.1.1.1
huawei(config)#mpls
huawei(config-mpls)#mpls te
huawei(config-mpls)#quit
```

     ii. Enable the basic MPLS and MPLS TE on the interface.

```
huawei(config)#vlan 10 standard
huawei(config)#mpls vlan 10
huawei(config)#port vlan 10 0/19 0
huawei(config)#interface vlanif 10
huawei(config-if-vlanif10)#ip address 10.1.2.10 24
huawei(config-if-vlanif10)#mpls
huawei(config-if-vlanif10)#mpls te
huawei(config-if-vlanif10)#quit
huawei(config)#vlan 21 standard
huawei(config)#mpls vlan 21
huawei(config)#port vlan 21 0/19 1
huawei(config)#interface vlanif 21
huawei(config-if-vlanif21)#ip address 10.1.1.10 24
huawei(config-if-vlanif21)#mpls
huawei(config-if-vlanif21)#mpls te
huawei(config-if-vlanif21)#quit
```

  c. Configure the MPLS TE tunnel from the source end to the destination end.

     Configure the MPLS TE tunnel bound to the detected LSP.

```
huawei(config)#interface tunnel 10
huawei(config-if-tunnel10)#tunnel-protocol mpls te
huawei(config-if-tunnel10)#destination 3.3.3.3
huawei(config-if-tunnel10)#mpls te tunnel-id 20
huawei(config-if-tunnel10)#mpls te signal-protocol static
huawei(config-if-tunnel10)#mpls te commit
huawei(config-if-tunnel10)#quit
```

  d. Configure the static LSP bound to the MPLS TE tunnel.

     Source end MA5600T/MA5603T/MA5608T functions as the ingress of the detected static LSP.

```
huawei(config)#static-lsp ingress tunnel-interface tunnel 10
destination 3.3.3.3 nexthop 10.1.2.20 out-label 8192
```

     Source end MA5600T/MA5603T/MA5608T functions as the egress of the detected static LSP.

```
huawei(config)#static-lsp egress LSP1 incoming-interface vlanif 10 in-label
8193
```

Source end MA5600T/MA5603T/MA5608T functions as the egress of the backward static LSP.

```
huawei(config)#static-lsp egress LSP2 incoming-interface vlanif 20 in-label 8201
```

e. Enable MPLS OAM at source end MA5600T/MA5603T/MA5608T_A.

```
huawei(config)#mpls
huawei(config-mpls)#mpls oam
huawei(config-mpls)#quit
huawei(config)#mpls oam ingress tunnel 10 type ffd frequency 100 backward-lsp lsr-id 3.3.3.3 tunnel-id 20
...//Configure the MPLS OAM source end. Configure the tunnel ID of the detected LSP to 10, detection packet type to FFD, Tx frequency to 100 ms,  LSR-ID of the backward LSP to 3.3.3.3,
...//and backward LSP tunnel ID to 20.
huawei(config)#mpls oam ingress enable all
```

f. Save the data.

```
huawei(config)#save
```

- **Configure Router A or Router B.**

  When functioning as the transit node, Router A or Router B mainly forwards MPLS labels. The ingress interface, in label, next hop IP address, and out label must be configured bi-directionally. For detailed configuration, see the configuration guide of the specific router.

- **Configure destination end MA5600T/MA5603T/MA5608T_B.**

  a. Configure the loopback interface.

```
huawei(config)#interface loopback 0
huawei(config-if-loopback0)#ip address 3.3.3.3 32
huawei(config-if-loopback0)#quit
```

  b. Enable the basic MPLS and MPLS TE.

  i. Enable the basic MPLS and MPLS TE globally.

```
huawei(config)#mpls lsr-id 3.3.3.3
huawei(config)#mpls
huawei(config-mpls)#mpls te
huawei(config-mpls)#quit
```

  ii. Enable the basic MPLS and MPLS TE on the interface.

```
huawei(config)#vlan 11 standard
huawei(config)#mpls vlan 11
huawei(config)#port vlan 11 0/19 0
huawei(config)#interface vlanif 11
huawei(config-if-vlanif11)#ip address 10.1.3.20 24
huawei(config-if-vlanif11)#mpls
huawei(config-if-vlanif11)#mpls te
huawei(config-if-vlanif11)#quit
huawei(config)#vlan 20 standard
huawei(config)#mpls vlan 20
huawei(config)#port vlan 20 0/19 1
huawei(config)#interface vlanif 20
huawei(config-if-vlanif20)#ip address 10.1.4.20 24
huawei(config-if-vlanif20)#mpls
```

```
huawei(config-if-vlanif20)#mpls te
huawei(config-if-vlanif20)#quit
```

c. Configure the MPLS TE tunnel from the destination end to the source end.

Configure the MPLS TE tunnel bound to the detected LSP.

```
huawei(config)#interface tunnel 10
huawei(config-if-tunnel10)#tunnel-protocol mpls te
huawei(config-if-tunnel10)#destination 1.1.1.1
huawei(config-if-tunnel10)#mpls te tunnel-id 10
huawei(config-if-tunnel10)#mpls te signal-protocol static
huawei(config-if-tunnel10)#mpls te commit
huawei(config-if-tunnel10)#quit
```

Configure the MPLS TE tunnel bound to the backward LSP.

```
huawei(config)#interface tunnel 20
huawei(config-if-tunnel20)#tunnel-protocol mpls te
huawei(config-if-tunnel20)#destination 1.1.1.1
huawei(config-if-tunnel20)#mpls te tunnel-id 20
huawei(config-if-tunnel20)#mpls te signal-protocol static
huawei(config-if-tunnel20)#mpls te commit
huawei(config-if-tunnel20)#quit
```

d. Configure the static LSP bound to the tunnel.

Destination end MA5600T/MA5603T/MA5608T functions as the egress of the detected static LSP.

```
huawei(config)#static-lsp egress LSP2 incoming-interface vlanif 10 in-label
8192
```

Destination end MA5600T/MA5603T/MA5608T functions as the ingress of the detected static LSP.

```
huawei(config)#static-lsp ingress tunnel-interface tunnel 10
destination 1.1.1.1 nexthop 10.1.3.10 out-label 8193
```

Destination end MA5600T/MA5603T/MA5608T functions as the ingress of the backward static LSP.

```
huawei(config)#static-lsp ingress tunnel-interface tunnel 20
destination 1.1.1.1 nexthop 10.1.4.10 out-label 8200
```

e. Enable MPLS OAM at destination end MA5600T/MA5603T/MA5608T.

```
huawei(config)#mpls
huawei(config-mpls)#mpls oam
huawei(config-mpls)#quit
huawei(config)#mpls oam egress lsr-id 1.1.1.1 tunnel-id 10 type ffd frequency
100 backward-lsp t
unnel 20 private
...//Configure the MPLS OAM destination end. Configure the ingress LSR-ID of
the detected LSP to 1.1.1.1, tunnel ID to 10, detection packet type to FFD,
Tx frequency to 100 ms,
...//backward LSP tunnel ID to 20, and tunnel to exclusive mode.
huawei(config)#mpls oam egress enable all
```

f. Save the data.

```
huawei(config)#save
```

----End

Result

After the configuration, shut down the interface of VLAN 10 by running the **shutdown** command on MA5600T/MA5603T/MA5608T_A to simulate the link fault:

- On MA5600T/MA5603T/MA5608T_B, run the **display mpls oam egress** command and you can see the following defect state: dLocv detected (dLocv).
- On MA5600T/MA5603T/MA5608T_A, run the **display mpls oam ingress** command and you can see the following defect state: in defect (In-defect).

Perform similar operations on MA5600T/MA5603T/MA5608T_B and you can obtain similar results.

Configuration Example of the MPLS OAM Protection Switching Function

This topic describes how to configure MPLS OAM to implement the protection switching function.

Service Requirements

- The OAM mechanism is used to detect in real time whether the MPLS link is normal and generates an alarm in time when a link fault is detected.
- The end-to-end tunnel protection technology is provided to recover the interrupted service.
- RSVP-TE is used to create an LSP tunnel for the specified path and reserve resources so that the existing bandwidth resources can be fully used and QoS can be improved for specific services.

Prerequisite

- The OSPF protocol must be successfully configured on each LSR in the network (the host route of each port must be successfully advertised).
- The interface IP address and mask, loopback interface, and LSR-ID must be configured on each LSR.
- The global and physical interface MPLS and MPLS TE functions must be enabled on each node of the LSR.

Networking

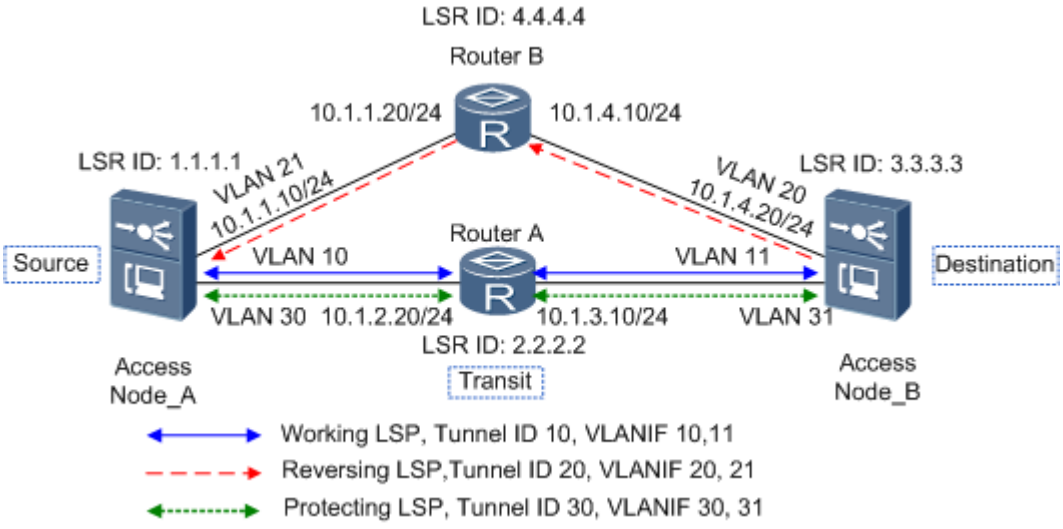
Figure 10-15 shows an example network for configuring the MPLS OAM protection switching function.

Configure two LSP tunnels on source end MA5600T/MA5603T/MA5608T_A and destination end MA5600T/MA5603T/MA5608T_B functioning primary and secondary LSPs. Enable the MPLS OAM protection switching function for the LSPs. When the primary LSP is faulty, the traffic is switched to the secondary LSP. Configure the backward LSP for reporting a fault to source end MA5600T/MA5603T/MA5608T_A.

NOTE

To prevent a fault from occurring on a transit node (for example, router A), it is recommended that you specify different transit nodes when creating a secondary LSP.

Figure 10-15 Configuring the MPLS OAM protection switching function



Data Plan

Table 10-6 provides the data plan for the MPLS OAM protection switching.

Table 10-6 Data plan for the MPLS OAM protection switching

Item	Data
MA5600T/MA5603T/MA5608T_A	LSR ID: 1.1.1.1
	Port: 0/19/0 IP address of VLAN interface 10 connected to Router A: 10.1.2.10/24
	Port: 0/19/1 IP address of VLAN interface 30 connected to Router A: 10.1.5.10/24 IP address of VLAN interface 21 connected to Router B: 10.1.1.10/24
MA5600T/MA5603T/MA5608T_B	LSR ID: 3.3.3.3
	Port: 0/19/0 IP address of VLAN interface 11 connected to Router A: 10.1.3.20/24
	Port: 0/19/1 IP address of VLAN interface 20 connected to Router B: 10.1.4.20/24 IP address of VLAN interface 31 connected to Router A: 10.1.6.20/24
	Backward tunnel: Router B to MA5600T/MA5603T/MA5608T_A

Item	Data
Router A	LSR ID: 2.2.2.2
Router B	LSR ID: 4.4.4.4

Procedure

- **Configure source end MA5600T/MA5603T/MA5608T_A.**
  a. Configure the loopback interface.

```
huawei(config)#interface loopback 0
huawei(config-if-loopback0)#ip address 1.1.1.1 32
huawei(config-if-loopback0)#quit
```

  b. Enable the basic MPLS, MPLS TE, and RSVP-TE functions.
    i. Enable the global basic MPLS, MPLS TE, and RSVP-TE functions.

```
huawei(config)#mpls lsr-id 1.1.1.1
huawei(config)#mpls
huawei(config-mpls)#mpls te
huawei(config-mpls)#mpls rsvp-te
huawei(config-mpls)#mpls te cspf
huawei(config-mpls)#quit
```

    ii. Enable the interface basic MPLS, MPLS TE, and RSVP-TE functions.

```
  //Configure the attributes of VLAN interface 10 and configure the IP
address of VLAN interface10 to 10.1.2.10/24.
huawei(config)#vlan 10 standard
huawei(config)#mpls vlan 10
huawei(config)#port vlan 10 0/19 0
huawei(config)#interface vlanif 10
huawei(config-if-vlanif10)#ip address 10.1.2.10 24
huawei(config-if-vlanif10)#mpls
huawei(config-if-vlanif10)#mpls te
huawei(config-if-vlanif10)#mpls rsvp-te
huawei(config-if-vlanif10)#mpls te bandwidth max-reservable-bandwidth
10240
  //(Optional) Configure VLAN interface 10 to provide a reservable
bandwidth of 10240 kbit/s for all tunnels.
huawei(config-if-vlanif10)#quit

  //Configure the attributes of VLAN interface 30 and configure the IP
address of VLAN interface 30 to 10.1.5.10/24.
huawei(config)#vlan 30 standard
huawei(config)#mpls vlan 30
huawei(config)#port vlan 30 0/19 1
huawei(config)#interface vlanif 30
huawei(config-if-vlanif30)#ip address 10.1.5.10 24
huawei(config-if-vlanif30)#mpls
huawei(config-if-vlanif30)#mpls te
huawei(config-if-vlanif30)#mpls rsvp-te
huawei(config-if-vlanif30)#mpls te bandwidth max-reservable-bandwidth
10240
```

```
  //(Optional) Configure VLAN interface 30 to provide a reservable
bandwidth of 10240 kbit/s for all tunnels.
huawei(config-if-vlanif30)#quit

  //Configure the attributes of VLAN interface 21 and configure the IP
address of VLAN interface 21 to 10.1.1.10/24.
huawei(config)#vlan 21 standard
huawei(config)#mpls vlan 21
huawei(config)#port vlan 21 0/19 1
huawei(config)#interface vlanif 21
huawei(config-if-vlanif21)#ip address 10.1.1.10 24
huawei(config-if-vlanif21)#mpls
huawei(config-if-vlanif21)#mpls te
huawei(config-if-vlanif21)#mpls rsvp-te
huawei(config-if-vlanif21)#mpls te bandwidth max-reservable-bandwidth
10240
  //(Optional) Configure VLAN interface 21 to provide a reservable
bandwidth of 10240 kbit/s for all tunnels.
huawei(config-if-vlanif21)#quit
```

c. Enable MPLS TE for the OSPF area.

```
huawei(config)#ospf 100
huawei(config-ospf-100)#opaque-capability enable
huawei(config-ospf-100)#area 0
huawei(config-ospf-100-area-0.0.0.0)#mpls-te enable standard-complying
huawei(config-ospf-100-area-0.0.0.0)#quit
huawei(config-ospf-100)#quit
```

d. Configure the MPLS TE tunnel from the source end to the destination end.

Configure the attributes of the working MPLS TE tunnel from the source end to the destination end.

```
huawei(config)#interface tunnel 10
huawei(config-if-tunnel10)#tunnel-protocol mpls te
huawei(config-if-tunnel10)#destination 3.3.3.3
huawei(config-if-tunnel10)#mpls te tunnel-id 10
huawei(config-if-tunnel10)#mpls te signal-protocol rsvp-te
huawei(config-if-tunnel10)#mpls te bandwidth ct0 5120   //(Optional) Configure
the global bandwidth of tunnel 10 to 5210 kbit/s.
huawei(config-if-tunnel10)#mpls te commit
huawei(config-if-tunnel10)#quit
```

Configure the attributes of the protection MPLS TE tunnel from the source end to the destination end.

```
huawei(config)#interface tunnel 30
huawei(config-if-tunnel30)#tunnel-protocol mpls te
huawei(config-if-tunnel30)#destination 3.3.3.3
huawei(config-if-tunnel30)#mpls te tunnel-id 30
huawei(config-if-tunnel30)#mpls te signal-protocol rsvp-te
huawei(config-if-tunnel30)#mpls te bandwidth ct0 5120   //(Optional) Configure
the global bandwidth of tunnel 30 to 5210 kbit/s.
huawei(config-if-tunnel30)#mpls te commit
huawei(config-if-tunnel30)#quit
```

e. Configure a tunnel protect group.

Configure tunnel 30 as the protect tunnel for tunnel 10, switching mode to revertive, and automatic WTR time to 900s (the corresponding WTR is 30 with step 30s).

```
huawei(config)#interface tunnel 10
huawei(config-if-tunnel10)#mpls te protection tunnel 30 mode revertive wtr 30
huawei(config-if-tunnel10)#mpls te commit
huawei(config-if-tunnel10)#quit
```

f. Enable MPLS OAM at source end MA5600T/MA5603T/MA5608T_A.

```
huawei(config)#mpls
huawei(config-mpls)#mpls oam
huawei(config-mpls)#quit
huawei(config)#mpls oam ingress tunnel 10 type ffd frequency 100
backward-lsp lsr-id 3.3.3.3 tunnel-id 20
  //Configure the MPLS OAM source end. Configure the tunnel ID of the detected
LSP to 10, detection packet type to FFD, Tx frequency to 100 ms, LSR-ID of the
backward LSP to 3.3.3.3,
  //and backward LSP tunnel ID to 20.
huawei(config)#mpls oam ingress enable all
```

g. Save the data.

```
huawei(config)#save
```

- **Configure Router A or Router B.**

  When functioning as the transit node, Router A or Router B mainly forwards MPLS labels. The ingress interface, in label, next hop IP address, and out label must be configured bi-directionally. For detailed configuration, see the configuration guide of the specific router.

- **Configure destination end MA5600T/MA5603T/MA5608T_B.**

  a. Configure the loopback interface.

```
huawei(config)#interface loopback 0
huawei(config-if-loopback0)#ip address 3.3.3.3 32
```

huawei(config-if-loopback0)#**quit**

  b. Enable the basic MPLS, MPLS TE, and RSVP-TE functions.

    i. Enable the global basic MPLS, MPLS TE, and RSVP-TE functions.

```
huawei(config)#mpls lsr-id 3.3.3.3
huawei(config)#mpls
huawei(config-mpls)#mpls te
huawei(config-mpls)#mpls rsvp-te
huawei(config-mpls)#mpls te cspf
huawei(config-mpls)#quit
```

    ii. Enable the interface basic MPLS, MPLS TE, and RSVP-TE functions.

```
  //Configure the attributes of VLAN interface 11 and configure the IP
address of VLAN interface 11 to 10.1.3.20/24.
huawei(config)#vlan 11 standard
huawei(config)#mpls vlan 11
huawei(config)#port vlan 11 0/19 0
huawei(config)#interface vlanif 11
huawei(config-if-vlanif11)#ip address 10.1.3.20 24
huawei(config-if-vlanif11)#mpls
huawei(config-if-vlanif11)#mpls te
```

```
huawei(config-if-vlanif11)#mpls rsvp-te
huawei(config-if-vlanif10)#quit

  //Configure the attributes of VLAN interface 20 and configure the IP address of VLAN interface 20 to 10.1.4.20/24.
huawei(config)#vlan 20 standard
huawei(config)#mpls vlan 20
huawei(config)#port vlan 20 0/19 1
huawei(config)#interface vlanif 20
huawei(config-if-vlanif20)#ip address 10.1.4.20 24
huawei(config-if-vlanif20)#mpls
huawei(config-if-vlanif20)#mpls te
huawei(config-if-vlanif20)#mpls rsvp-te
huawei(config-if-vlanif20)#quit

  //Configure the attributes of VLAN interface 31 and configure the IP address of VLAN interface 31 to 10.1.6.20/24.
huawei(config)#vlan 31 standard
huawei(config)#mpls vlan 31
huawei(config)#port vlan 31 0/19 1
huawei(config)#interface vlanif 31
huawei(config-if-vlanif31)#ip address 10.1.6.20 24
huawei(config-if-vlanif31)#mpls
huawei(config-if-vlanif31)#mpls te
huawei(config-if-vlanif31)#mpls rsvp-te
huawei(config-if-vlanif31)#quit
```

c. Configure the MPLS TE tunnel bound to the backward LSP.

Configure the tunnel ID to 20, destination IP address to 1.1.1.1, and global bandwidth for the tunnel to 5120 kbit/s.

```
huawei(config)#interface tunnel 20
huawei(config-if-tunnel20)#tunnel-protocol mpls te
huawei(config-if-tunnel20)#destination 1.1.1.1
huawei(config-if-tunnel20)#mpls te tunnel-id 20
huawei(config-if-tunnel20)#mpls te signal-protocol rsvp-te
huawei(config-if-tunnel20)#mpls te bandwidth ct0 5120
huawei(config-if-tunnel20)#mpls te reserved-for-binding
huawei(config-if-tunnel20)#mpls te commit
huawei(config-if-tunnel20)#quit
```

d. Enable MPLS OAM at destination end MA5600T/MA5603T/MA5608T_B.

```
huawei(config)#mpls
huawei(config-mpls)#mpls oam
huawei(config-mpls)#quit
huawei(config)#mpls oam egress lsr-id 1.1.1.1 tunnel-id 10 type ffd frequency 100
 backward-lsp tunnel 20 private
  //Configure the MPLS OAM destination end. Configure the ingress LSR-ID of the detected LSP to 1.1.1.1, tunnel ID to 10, detection packet type to FFD, Tx frequency to 100 ms,
  //backward LSP tunnel ID to 20, and tunnel to exclusive mode.
huawei(config)#mpls oam egress enable all
```

e. Save the data.

```
huawei(config)#save
```

----End

Result

After the configuration, you can shut down the interface of VLAN 10 by running the **shutdown** command on MA5600T/MA5603T/MA5608T_A to simulate the link fault. Then, you can query the information about the primary tunnel (with ID 10) that is configured on MA5600T/MA5603T/MA5608T_A by running the **display mpls te protection tunnel** command on MA5600T/MA5603T/MA5608T_A. The information is as follows:

- Status of the working tunnel (work-tunnel defect state): in defect.
- Status of the protection tunnel (protect-tunnel defect state): non-defect.
- Switch result: The traffic is switched to protection tunnel 30.

11 VPLS

About This Chapter

The Virtual Private LAN Service (VPLS), also called the Transparent LAN Service (TLS) or virtual private switched network service, is a Layer 2 VPN (L2VPN) technology that is based on Multi-Protocol Label Switching (MPLS) and Ethernet technologies.

11.1 What Is VPLS

Definition

The Virtual Private LAN Service (VPLS), also called the Transparent LAN Service (TLS) or virtual private switched network service, is a Layer 2 VPN (L2VPN) technology that is based on Multi-Protocol Label Switching (MPLS) and Ethernet technologies.

Purpose

The primary goal of VPLS is to interconnect multiple Ethernet LANs through the Packet Switched Network (PSN). In this manner, these LANs can function as one LAN. VPLS can implement the multipoint-to-multipoint VPN networking; therefore, by using the VPLS technology, service providers (SPs) can provide the Ethernet-based multipoint services through MPLS backbone networks. In addition, by utilizing the VPLS solution in which MPLS virtual circuits (VCs) function as the Ethernet bridge links, SPs can transparently transmit LAN services on the MPLS network.

11.2 References

The following table lists the references of this document.

Document No.	Description
RFC 4762	Virtual Private LAN Service (VPLS) Using Label Distribution Protocol (LDP) Signaling
draft-ietf-l2vpn-oam-req-frmk-01	VPLS OAM Requirements and Framework

11.3 Principles

11.3.1 VPLS Introduction

Basic VPLS Transport Structure

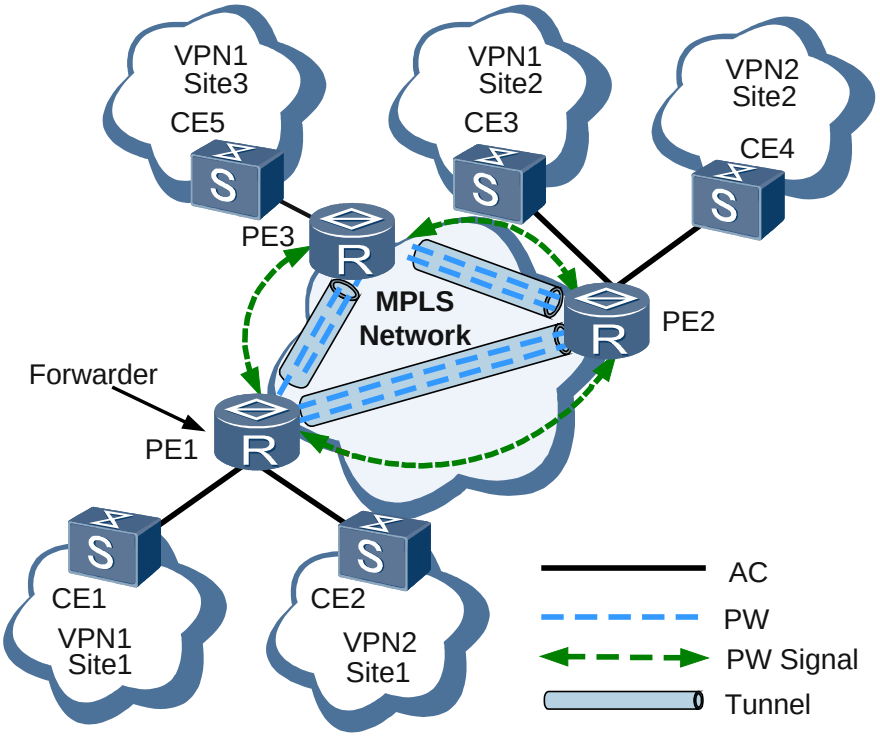
Figure 11-1 shows an example of a VPLS network. The entire VPLS network is similar to a switch. PWs are established over MPLS tunnels between VPN sites to transparently transmit Layer 2 packets between sites. When forwarding packets, PEs learn the source MAC addresses of these packets and create MAC entries, mapping MAC addresses to attachment circuits (ACs) and PWs.

The following table describes the various concepts related to VPLS networks.

Table 11-1 Description of VPLS concepts

Name	Description
AC	A link between a CE and a PE. An AC must be established using Ethernet interfaces. On a VPLS network, AC interfaces can be Ethernet interfaces, Ethernet sub-interfaces, VLANIF interfaces, Eth-Trunk interfaces, Eth-Trunk sub-interfaces, VE interfaces, QinQ interfaces, and VE (ATM 1483B) interfaces.
PW	A bidirectional virtual connection between two virtual switch instances (VSIs) residing on two PEs. A PW consists of a pair of unidirectional MPLS VCs transmitting in opposite directions.
VSI	A type of instance used to map ACs to PWs. A VSI independently provides VPLS services and forwards Layer 2 packets based on MAC addresses and VLAN tags. A VSI has the Ethernet bridge function and can terminate PWs.
PW signaling	A type of signaling used to create and maintain PWs. PW signaling is the foundation for VPLS implementation. Currently, the PW signaling is LDP or BGP. MA5600T/MA5603T/MA5608T supports only LDP PW signaling.
Tunnel	A connection between a local PE and a remote PE used to transparently transmit data between PEs. A tunnel can carry multiple PWs. MA5600T/MA5603T/MA5608T supports only MPLS tunnels.
Forwarder	Similar to a VPLS forwarding table. After a PE receives packets from an AC, the forwarder of the PE selects a PW to forward these packets.

Figure 11-1 Basic VPLS transmission process



The forwarding of a packet from CE1 to CE3 on VPN1 is used as an example:

1. CE1 sends a Layer 2 packet to PE1 over an AC.
2. After PE1 receives the packet, the forwarder of PE1 selects a PW for forwarding the packet.
3. PE1 then adds two MPLS labels to the packet based on the PW forwarding entry and sends the packet to PE2. The private network label identifies the PW, and the public network label identifies the tunnel between PE1 and PE2.
4. After PE2 receives the packet from the public tunnel, PE2 removes the private network label of the packet.
5. The forwarder of PE2 selects an AC and forwards the packet to CE3 over the AC.

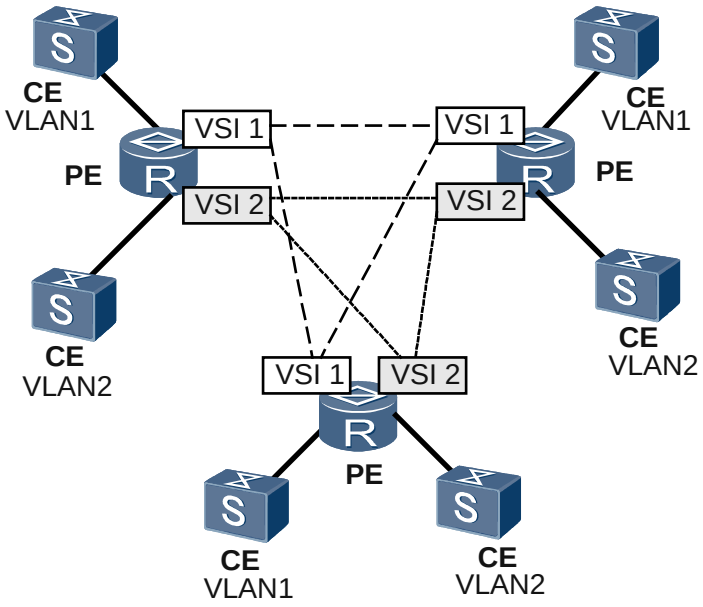
VPLS Implementation Process

Transmission of packets between CEs relies on VSIs configured on PEs, and PWs established between the VSIs. Figure 11-2 shows transmission of Ethernet frames over full-mesh PWs between PEs.

The Ethernet often uses the Spanning Tree Protocol (STP) to prevent loops. VPLS networks, however, use full-mesh PWs and split horizon to avoid loops as follows:

- The PEs in a VSI must be fully meshed. That is, a PE must create a tree path to every other PE in the VSI.
- Each PE must support split horizon to avoid loops. Split horizon requires that packets received from a PW in a VSI should not be forwarded to other PWs in the VSI. Any two PEs in a VSI must communicate over a direct PW, which is why full-mesh PWs are required between PEs in a VSI.

Figure 11-2 VPLS forwarding model



A VPLS network consists of a control plane and a forwarding plane.

- The control plane of a VPLS PE provides the following functions:
  - Member discovery: a process in which a PE in a VSI discovers the other PEs in the same VSI. This process can be implemented manually or automatically using protocols. BGP VPLS and BGP AD VPLS both support automatic member discovery.
  - Signaling mechanism: PWs between PEs in the same VSI are established, maintained, or torn down using signaling protocols such as LDP and BGP.
- The forwarding plane of a VPLS PE provides the following functions:
  - Encapsulation: After receiving Ethernet frames from a CE, a PE encapsulates the frames into packets and sends the packets to a PSN.
  - Forwarding: A PE determines how to forward a packet based on the inbound interface and destination MAC address of the packet.
  - Decapsulation: After receiving packets from a PSN, a PE decapsulates these packets into Ethernet frames and sends the frames to a CE.

VPLS Implementation Modes

VPLS can be implemented in LDP, BGP, or BGP AD mode.

- VPLS implemented in LDP mode is also called Martini VPLS.
- VPLS implemented in BGP mode is also called Kompella VPLS.
- VPLS BGP AD uses extended BGP Update packets to implement automatic member discovery. It also uses LDP FEC 129 signaling packets for local and remote VSIs to automatically negotiate and establish VPLS PWs.

The differences between the three tunnel setup modes are as follows:

- In LDP tunnel setup mode, the requirements for PEs are low, but no auto-discovery mechanism for VPN members can be provided, which has to be configured manually. In

BGP tunnel setup mode, the requirements for PEs are high. That is, PEs must run BGP. In addition, the auto-discovery mechanism for VPN members can be provided.

- In LDP tunnel setup mode, an LDP session must be created between every two PEs. The number of sessions is in direct ratio to the square of the number of PEs. In BGP tunnel setup mode, route reflector (RR) can be used to reduce the number of BGP connections.
- In LDP tunnel setup mode, each PE is assigned with a label only if necessary. In BGP tunnel setup mode, each PE is assigned with a label block, which leads to the waste of labels.
- In LDP tunnel setup mode, the VSIs configured in all domains must use the same VSI ID range. In BGP tunnel setup mode, the VPN target is used to identify VPNs.

Table 11-2 shows the comparison between the two VPLS tunnel setup modes.

Table 11-2 Comparison between two VPLS tunnel setup modes

Type	LDP	BGP
Requirements for PEs	Common	High
Auto-discovery supported	No	Yes
Implementation complexity	Low	High
Expansibility	Poor	Good
Label utilization ratio	High	Low
Configuration workload	High	Low
Cross-domain restrictions	High	Low

After the preceding comparison, the following conclusions can be drawn:

- The LDP tunnel setup mode is preferable when the number of VPLS sites is relatively small, the VPLS network seldom or never traverses multiple domains, and PEs do not run BGP.
- The BGP tunnel setup mode is applicable at the core layer of a large-scale network when PEs run BGP and cross-domain is required.

If the scale of a VPLS network is large (a great number of nodes in a wide geographical range), you can use HVPLS to combine the two modes. That is, the core layer uses the BGP tunnel setup mode and the access layer uses the LDP tunnel setup mode.

VPLS assumes that each PE is capable of setting up tunnels; PW labels functions as the identifiers for services; tunnels are responsible for transmitting VPLS data from a PE to another PE.

VPLS Encapsulation Modes

- Packet encapsulation on ACs

  Packet encapsulation on ACs depends on the user access mode, which can be VLAN or Ethernet access.

NOTE

Currently, the MA5600T/MA5603T/MA5608T supports only packet encapsulation type of VLAN.

Table 11-3 Packet encapsulation on ACs

Packet Encapsulation Type	Description
VLAN	The header of each Ethernet frame sent between CEs and PEs carries a VLAN tag, known as the SVLAN. This is a service delimiter identifying users on an ISP network.
Ethernet	The header of each Ethernet frame sent between CEs and PEs does not carry a SVLAN. If the frame header contains a VLAN tag, it is an inner VLAN tag called the CVLAN. A CE does not add the CVLAN to an Ethernet frame; instead, the tag is carried in a packet before the packet is sent to the CE. A CVLAN informs the CE to which VLAN the packet belongs, and is meaningless to PEs.

- Packet encapsulation on PWs

  The PW ID and PW encapsulation type uniquely identify a PW. The PW IDs and PW encapsulation types configured on the two end PEs of a PW must be the same. The packet encapsulation types of packets on PWs can be raw or tagged. By default, packets on PWs are encapsulated in tagged mode.

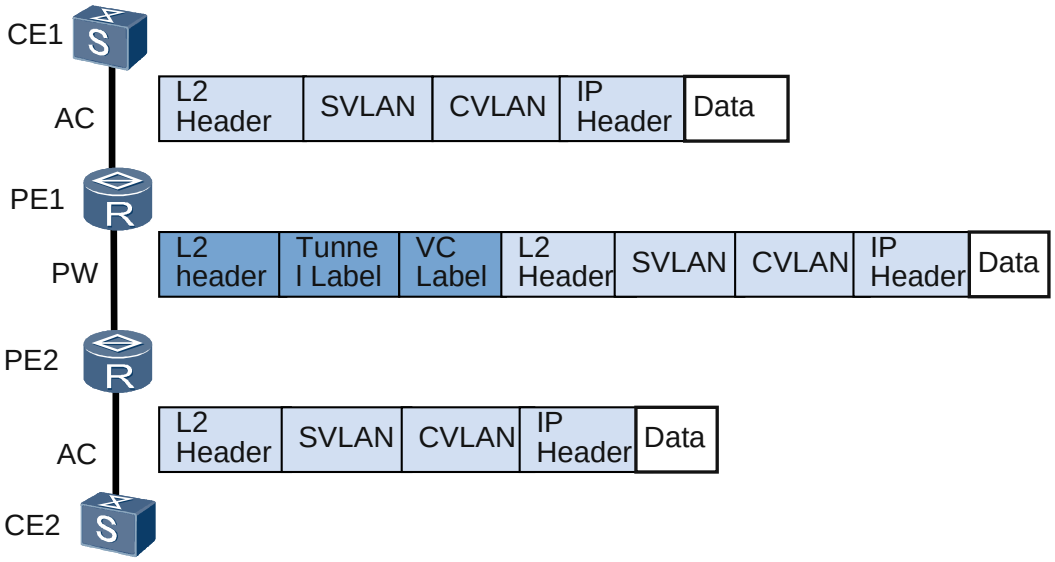
Table 11-4 Packet encapsulation on PWs

Packet Encapsulation Type	Description
Raw	Packets transmitted over a PW cannot carry SVLANs. If a PE receives a packet with the SVLAN from a CE, the PE strips the SVLAN and adds double labels (outer tunnel label and inner VC label) to the packet before forwarding it. If a PE receives a packet with no SVLAN from a CE, the PE directly adds double labels (outer tunnel label and inner VC label) to the packet before forwarding it. The PE determines whether to add the SVLAN to a packet based on actual configurations before sending it to a CE. The PE is not allowed to rewrite or remove an existing CVLAN.
Tagged	Packets transmitted over a PW must carry SVLANs. If a PE receives a packet with the SVLAN from a CE, the PE directly adds double labels (outer tunnel label and inner VC label) to the packet before forwarding it. If a PE receives a packet with no SVLAN from a CE, the PE adds a null SVLAN and double labels (outer tunnel label and inner VC label) to the packet before forwarding it. The PE determines whether to rewrite, remove, or preserve the SVLAN of a packet based on actual configurations before forwarding it to a CE.

Encapsulation modes of packets transmitted over ACs and PWs can be used together. The following uses VLAN+tagged encapsulation (with the CVLAN) as examples to describe the packet exchange process.

VLAN+tagged encapsulation (with the CVLAN)

Figure 11-3 VLAN+tagged encapsulation (with the CVLAN)



As shown in Figure 11-3, ACs use VLAN encapsulation and PWs use tagged encapsulation; packets transmitted from CEs to PEs carry U-Tags and SVLANs.

The packet exchange process is as follows:

1. CE1 sends a packet that has Layer 2 encapsulation and carries both a CVLAN and a SVLAN to PE1.
2. Upon receipt, PE1 does not process the two tags (PE1 retains the CVLAN because it treats the U-tag user data; PE1 retains the SVLAN because a packet sent to a PW with the tagged packet encapsulation mode must carry a SVLAN). PE1 searches the corresponding VSI for a forwarding entry and selects a tunnel and a PW to forward the packet based on the found forwarding entry. PE1 adds double labels (outer tunnel label and inner VC label) to the packet based on the selected tunnel and PW, performs Layer 2 encapsulation, and forwards the packet to PE2.
3. Upon receipt, PE2 removes the Layer 2 encapsulation carried out by PE1 and its double labels (outer tunnel label and inner VC label), and sends the original Layer 2 packet that carries the CVLAN and SVLAN to CE2.

The processing of sending a packet from CE2 to CE1 is similar to this process.

Derivative VPLS Functions

Traffic Statistics

Traffic statistics can be collected based on ACs or PWs, and the status of various types of traffic can be viewed in real time.

VPLS Service Isolation

VPLS service isolation allows you to prohibit communication between users that use the same service and bound to the same VSI.

By default, traffic can be forwarded between AC interfaces, between UPE PWs, and between AC interfaces and UPE PWs in a VSI. On a non-hierarchical VPLS network, VPLS service isolation prohibits traffic forwarding between AC interfaces. On an HVPLS network, VPLS

service isolation prohibits traffic forwarding between AC interfaces, between UPE PWs, and between AC interfaces and UPE PWs.

11.3.2 VPLS Layer 2 Functions

Background

A characteristic of the Ethernet is that a port sends unicast packets with unknown destination MAC addresses, broadcast packets, and multicast packets to all other ports on the Ethernet. As an Ethernet-based technology, VPLS emulates an Ethernet bridge for user networks. To forward packets on a VPLS network, PEs must establish MAC address tables and forward packets based on MAC addresses or MAC addresses and VLAN tags.

Related Concepts

- MAC address learning

  Table 11-5 describes MAC address learning modes.

Table 11-5 MAC address learning modes

MAC Address Learning Mode	Description	Characteristic
Qualified	A PE learns the MAC addresses and VLAN tags of received Ethernet frames. In this mode, each user VLAN is an independent broadcast domain and has independent MAC address space.	The broadcast domain is confined to each user VLAN. Qualified learning can result in large FIB table sizes, because the logical MAC address is now a VLAN tag + MAC address.
Unqualified	A PE learns only the MAC addresses of Ethernet frames. In this mode, all user VLANs share the same broadcast domain and MAC address space. The MAC address of each user VLAN must be unique.	If an AC interface is associated with multiple user VLANs, this AC interface must be a physical interface bound to a unique VSI.

NOTE

At present, the MA5600T/MA5603T/MA5608T supports only MAC address learning in qualified mode.

- MAC address aging

  An aging mechanism removes MAC entries that a PE no longer needs. If a MAC entry is not updated within a specified period of time, this entry will be aged.

Implementation

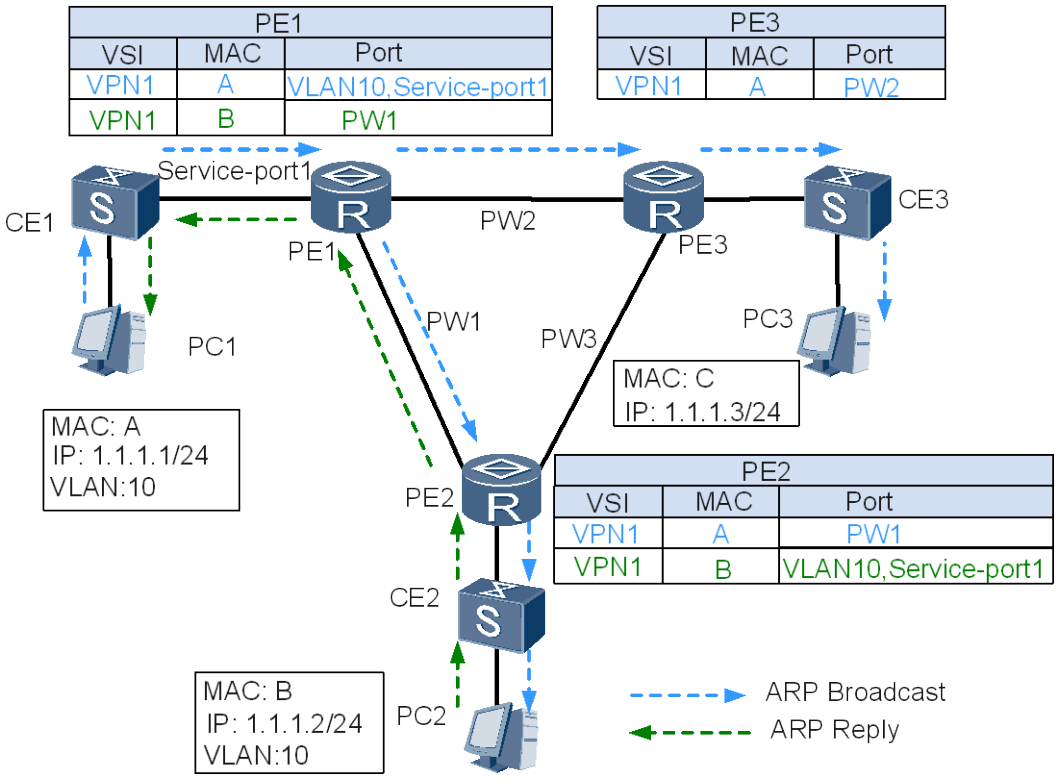
PEs establish MAC address tables based on dynamic MAC address learning and associates destination MAC addresses with PWs. Table 11-6 describes the MAC address learning process.

Table 11-6 MAC address learning process

MAC Address Learning Process	Description
Learning MAC addresses from user-side packets	After receiving packets from a CE, a PE maps their source MAC addresses to the service port corresponding to CE.
Learning MAC addresses from PW-side packets	A PW consists of a pair of MPLS VCs transmitting in opposite directions. A PW will go Up only after the two MPLS VCs are established. After a PE receives a packet with an unknown source MAC address from a PW, the PE maps the source MAC address to the AC interface receiving the packet.

Figure 11-4 shows the process of MAC address learning and flooding on a PE. PC1 and PC2 both belong to VLAN10. When PC1 pings IP address 1.1.1.2, PC1 does not know the MAC address corresponding to this IP address and advertises an ARP Request packet.

Figure 11-4 MAC address learning process



1. After receiving the ARP Request packet sent by PC1 from service port1 that connects to CE1, PE1 adds the MAC address of PC1 to its own MAC address table, as shown in the blue section of the MAC entry.
2. PE1 advertises the ARP Request packet to its other ports (PW1 and PW2 can be viewed as ports).

3. After receiving the ARP Request packet from PW1, PE2 adds the MAC address of PC1 to its own MAC address table, as shown in the blue section of the MAC entry.
4. Based on split horizon, PE2 sends the ARP Request packet to only the port connecting to CE2 (as indicated by the blue dashed line), but not to PW1. This ensures that only PC2 receives the ARP Request packet. VPLS split horizon ensures that packets received from public network PWs are forwarded to only private networks, not to other public network PWs.
5. After PC2 receives the ARP Request packet and finds that it is the destination of this packet, PC2 sends an ARP Reply packet to PC1 (as indicated by the green dashed line).
6. After receiving the ARP Reply packet from PC2, PE2 adds the MAC address of PC2 to its own MAC address table, as indicated by the blue section of the MAC entry. The destination MAC address of the ARP Reply packet is the MAC address of PC1 (MAC A). After searching its MAC address table, PE2 sends the ARP Reply packet to PE1 over PW1.
7. After receiving the ARP Reply packet from PE2, PE1 adds the MAC address of PC2 to its own MAC address table. After searching its MAC address table, PE1 sends the ARP Reply packet to PC1 through service port1.
8. After receiving the ARP Reply packet from PC2, PC1 completes MAC address learning.
9. While advertising the ARP Request packet to PW1, PE1 also advertises the ARP Request packet to PE3 over PW2. After receiving the ARP Request packet, PE3 adds the MAC address of PC1 to its MAC address table. Based on split horizon, PE3 sends the ARP Request packet to only PC3. Because PC3 is not the destination of the ARP Request packet, PC3 does not send any ARP Reply packet.

Derivative Functions

Traffic Restriction

On a VPLS network, you can limit the rates of broadcast, multicast, and unknown unicast packets to:

- Enhance traffic management and appropriately allocate user bandwidth.
- Prevent traffic attacks and enhance network security.

Limit on the Number of Learned MAC Addresses

After the number of MAC entries or MAC address learning time reaches the set threshold, a device forwards or drops newly received packets and decides whether to report an alarm to the network management system (NMS).

This function applies to networks with relatively fixed users but insufficient security, such as residential access networks and enterprise intranets without security management.

11.3.3 LDP VPLS

Background

LDP VPLS (Martini VPLS) uses a static discovery mechanism to discover VPLS members using LDP signaling. VPLS information is carried in extended TLV fields of LDP signaling packets.

Related Concepts

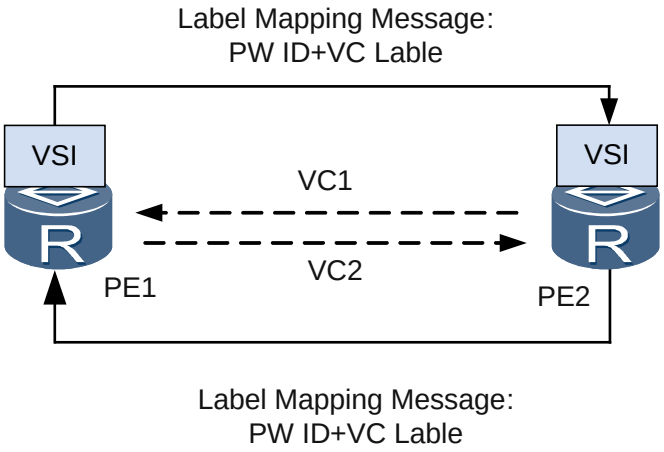
LDP VPLS involves the following concepts:

- FEC: A set of packets with similar or identical characteristics and forwarded in the same way by LSRs. Characteristics determining the FEC of a packet include the destination address, service type, and QoS attribute. Currently, the MA5600T/MA5603T/MA5608T only supports VLAN as FEC.
- TLV: A highly efficient and expansible coding mode for protocol packets. To support new features, you only need to add new types of TLVs to carry information required by the features.
- DU: A label distribution mode in which an LSR distributes labels to FECs without having to receive Label Request messages from its upstream LSR.
- Liberal: A label retention mode in which an LSR retains the label mapping received from a neighboring LSR, regardless of whether the neighboring LSR is its next hop. In liberal label retention mode, an LSR can use the labels sent from neighboring LSRs that are not at the next hop to re-establish an LSP. This mode requires more memory and label space than the conservative mode.

Implementation Process

- Figure 11-5 shows the process of establishing a PW using LDP signaling.

Figure 11-5 Establishing a PW using LDP signaling

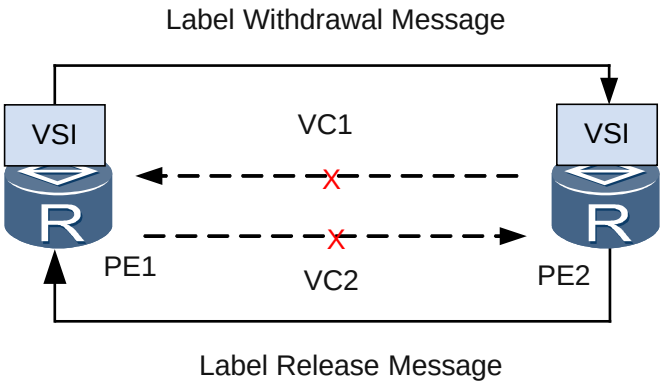


a. After PE1 is associated with a VSI, and PE2 is configured as a peer of PE1, PE1 sends a Label Mapping message to PE2 in DU mode if an LDP session already exists between PE1 and PE2. The Label Mapping message carries information required to establish a PW, such as the PW ID, VC label, and interface parameters.

b. Upon receipt of the message, PE2 checks whether itself has been associated with the VSI. If PE2 has been associated with the VSI and PW parameters on PE1 and PE2 are consistent, PE1 and PE2 belong to the same VSI. In this case, PE2 establishes a unidirectional VC named VC1 immediately after PE2 receives the Label Mapping message. Meanwhile, PE2 sends a Label Mapping message to PE1. After receiving the message, PE1 takes a similar sequence of actions to PE2 and establishes VC2.

- Figure 11-6 shows the process of tearing down a PW using LDP signaling.

Figure 11-6 Tearing down a PW using LDP signaling



a. After the peer configuration about PE2 is deleted from PE1, PE1 sends a Label Withdrawal message to PE2. After receiving the Label Withdrawal message, PE2 withdraws its local VC label, tears down VC1, and sends a Label Release message to PE1.

b. After receiving the Label Release message, PE1 withdraws its local VC label and tears down VC2.

Derivative Functions

MAC Withdrawal

- After receiving a MAC-Withdraw message that carries the NULL MAC TLV, the remote PE clears all MAC address entries in the VSI by default. You can configure a PE to delete MAC address entries in standard mode defined in RFC 4762. In standard mode, only MAC address entries for those ports that are not used by the corresponding PW are deleted.
- After receiving a MAC-Withdraw message that carries the PE-ID TLV, the remote PE clears the MAC address entry for the corresponding PW.

Ignorance of the AC Status by a VSI

Before the replacement of CEs, you can configure VSIs on UPEs to temporarily ignore the AC interface status check. Then, check whether VSIs on UPEs can work properly after new CEs are deployed. A VSI can be Up only if at least one AC interface and one PW is Up. After you configure a VSI to ignore the AC interface status check, the VSI remains Up as long as one PW is Up, regardless of whether the AC interface status is Up or Down.

Receiving of Group Messages by PWs

The IETF defines the usage scenario of this function. If multiple PWs, belonging to the same group and having the same status, are configured on a physical interface, Group messages can be used to notify PWs of the interface status change when the physical interface goes Up or Down, reducing the number of Notification messages required.

NOTE

At present, the MA5600T/MA5603T/MA5608T can only receive group messages and cannot send group messages.

PW Reliability

LDP VPLS ensures PW reliability by manual configuration. When the primary PW fails, traffic from the primary PW switch to the secondary PW; When primary PW recovers, traffic can be immediate or delayed switch back to the primary PW.

Usage Scenario

The LDP mode applies to VPLS networks that do not have many sites, do not span multiple ASs, or with PEs that do not run BGP.

Benefits

LDP VPLS brings the following benefits:

- Easy configuration
- Label resource saving

11.3.4 VPLS PW Redundancy

Implementation

To ensure the same forwarding capability, the PW redundancy protection mechanism to be used must allow the configuration of a single PW in a PW group to be an active PW and the remaining to be standby PWs, which requires corresponding signaling control.

RFC 4447 (Pseudowire Setup and Maintenance Using the Label Distribution Protocol [LDP]) specifies the PW Status TLV to transmit the PW forwarding status. The PW Status TLV is transported to the remote PW peer using a Label Mapping or LDP Notification message. The PW Status TLV is a 32-bit status code field. Each bit in the status code field can be set individually to indicate more than one failure. PW redundancy introduces a new PW status code 0x00000020. When the code is set, it indicates "PW forwarding standby".

Forwarding priorities (Primary or Secondary) must be configured for PWs that back up each other. The highest priority PW will be selected as the primary PW to forward traffic. The remaining PWs will be in the Secondary state to protect the primary PW.

NOTE
Currently, only one secondary PW can be configured for a primary PW.

The forwarding status of a PW determines whether the PW is used to forward traffic. The PW forwarding statuses depend on:

- Local and remote PW signaling statuses: A PE monitors the local signaling status and uses PW redundancy signaling to obtain remote signaling status from a remote PE.
- PW redundancy mode: Master/Slave or Independent mode is specified on PE1.
- PW forwarding priorities: PW forwarding priorities (Primary or Secondary) are specified on PE1.

Figure 11-7 shows that VPLS PW redundancy is configured on PE1. In normal cases, all local and remote PW signaling statuses on PE1 are Up. PEs at the two ends of a PW in different VPLS PW redundancy modes use different methods to select the same PW for transmitting user packets.

- In Master/Slave mode, PE1 determines local PW forwarding statuses based on preset forwarding priorities and inform PE2 and PE5 of the PW forwarding statuses; PE2 and PE5 determine their PW forwarding statuses based on the received PW primary and secondary statuses.

- In Independent mode, PE1 determines local PW forwarding statuses based on the forwarding statuses learned from PE2 and PE5; PE2 and PE5 determine their PW primary and secondary statuses based on signaling, which can be enhanced trunk (E-Trunk), enhanced automatic protection switching (E-APS), or Virtual Router Redundancy Protocol (VRRP) signaling, and notify PE1 of the forwarding statuses.

In both Master/Slave and Independent modes, if a primary PW is faulty, it becomes inactive and its secondary PW becomes active. PW-side faults do not affect the AC status. If AC-side faults occur (for example, a PE or AC link is faulty), the PW primary and secondary statuses in Independent mode will change because the statuses are determined by the master and backup statuses of the dual-homing devices; the PW primary and secondary statuses in Master/Slave mode will not change because they are determined by PW side.

NOTE

VPLS PW redundancy is similar to VPWS PW redundancy, with the exception that a virtual switch instance (VSI) has multiple PWs to different PEs. These PWs form various PW groups. PW switching in one group does not affect other PW groups.

Derivative Function

In addition to protection against network faults in real time, VPLS PW redundancy allows users to manually switch traffic between PWs in a group during network operation and maintenance. For example, if a device providing a primary PW needs to be maintained, a user can switch traffic to the secondary PW and switch it back to the primary PW after the maintenance.

NOTE

The interval between a switchover and a switchback must be at least 15s.

Usage Scenarios

VPLS PW redundancy can be used on hierarchical virtual private LAN service (HVPLS) networks and VPLS and virtual leased line (VLL) interconnected networks. These two types of networks can bear any services, but when newly planned or deployed, these networks are suggested to carry different services based on their networking characteristics.

- HVPLS networks are suitable for bearing multicast services, such as Internet Protocol television (IPTV) services, because HVPLS networks can save VPLS core network bandwidth. For details, see 11.4.3 VPLS PW Redundancy for Protecting Multicast Services.
- VPLS and VLL interconnected networks are suitable for bearing unicast services, such as high-speed internet (HSI) and voice over IP (VoIP) services, because VLL PEs do not need to learn user MAC addresses. For details, see 11.4.4 VPLS PW Redundancy for Protecting Unicast Services.

VPLS PW redundancy can also be used to improve reliability of existing networks. On the VPLS network in Figure 11-7, CE1 communicates with CE2, CE3, and CE4 through PWs between one VSI on PE1 and PE2, PE3, and PE4.

As services develop, services between CE1 and CE2, and between CE1 and CE3 require high reliability. Services between CE1 and CE4 do not require high reliability.

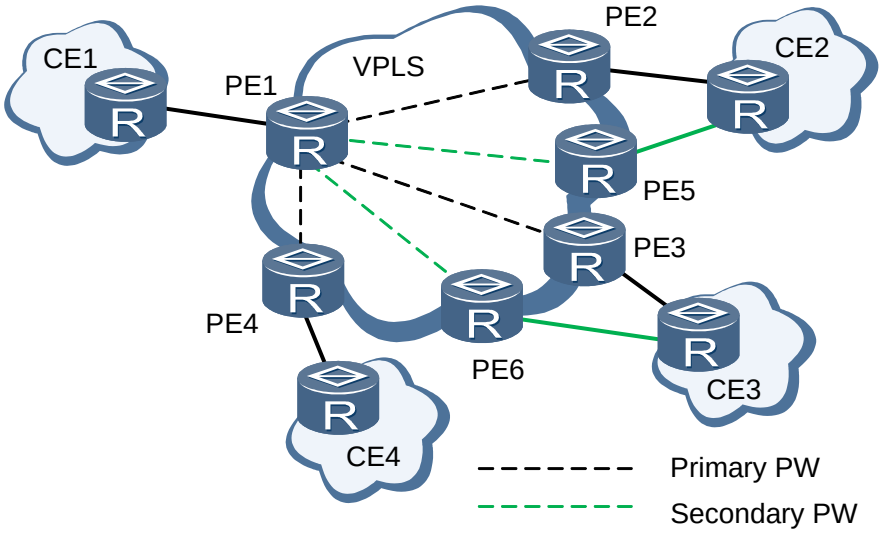
To meet the reliability requirements, PE5 and PE6 are deployed on the VPLS network to provide VPLS PW redundancy protection for PE2 and PE3, respectively. In addition, multiple PW groups to peer PEs are configured in one VSI on PE1. Links between CE1 and CE4 remain unchanged.

VPLS PW redundancy protects services against failures on the network side, AC side, or PEs without affecting existing services, improving network reliability.

NOTE

VPLS PW redundancy can be provided for the desired services without affecting services on other PWs, which reduces costs and maximizes profits.

Figure 11-7 VPLS PW redundancy networking



11.4 VPLS PW Redundancy Applications

11.4.1 Application of VPLS Individual Access

Service Overview

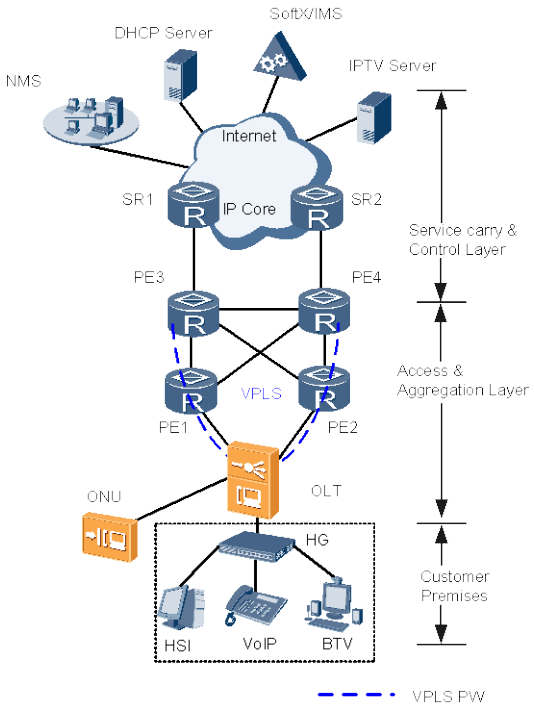
The traffic of individual services such as high speed internet (HSI), voice over IP (VoIP) and broadband TV (BTV) are carried by the carrier's metropolitan area network (MAN).

The traditional bearing technologies such as the asynchronous transfer mode (ATM) and frame relay (FR) have some defects such as high cost for network construction, slow speed and complicated deployment. Moreover, the traditional bearing technologies only support the point-to-point (P2P) interconnection for users. With the development of IP technology, the Ethernet-based virtual private LAN service (VPLS) technology supports transparent transmission of the above-mentioned individual services and achieves the point-to-multipoint (P2MP) interconnection for users. In addition, the Ethernet-based VPLS has many advantages, such as low cost for network construction, high speed and simple deployment. Therefore, the VPLS technology is widely used in the current MAN to transmit the user traffic.

Example Network

Figure 11-8 shows the VPLS individual access service.

Figure 11-8 Example network of VPLS individual access



The HSI service is used as an example in the example network.

- The MSAN/OLT is dual-homed to two AGS devices through the VPLS.
- The user HSI access service is provided through the PPPoE dialup and maps to the VPLS domain through a VLAN in upstream direction.
- PADI packets initiated from the user side are broadcast in the VPLS domain to which the packets belong. The broadcast packets are received on PE1 and PE2.
- The delay response is used between PE devices to terminate the dialups of some users so that the load sharing can be achieved.
- The split horizon between the VPLS and PW is enabled.

11.4.2 Application of VPLS Enterprise Access

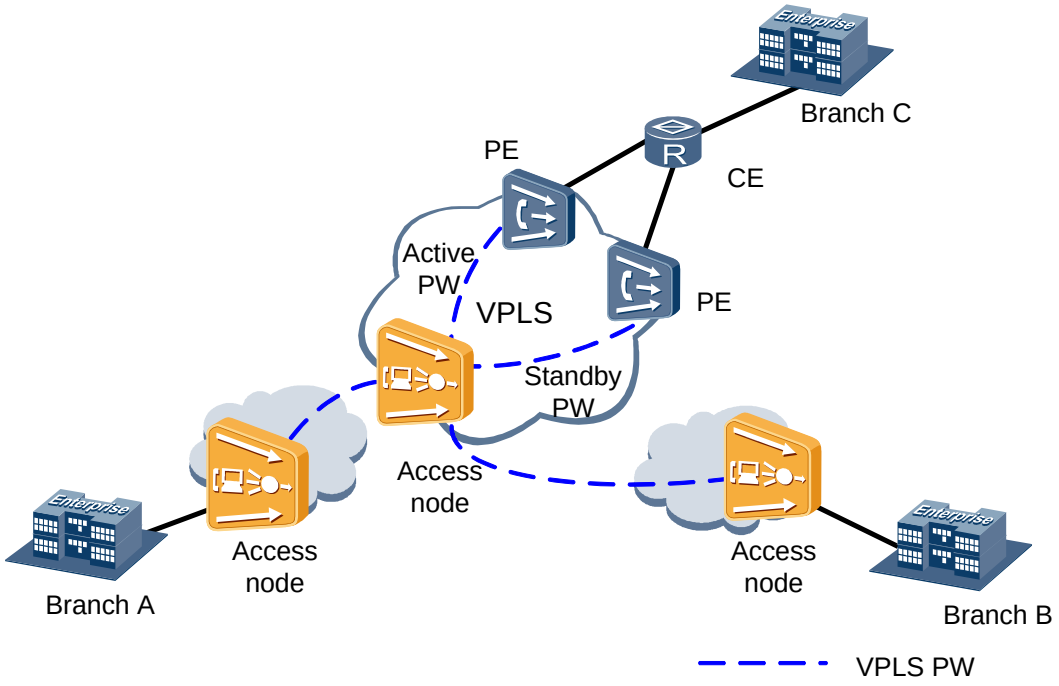
Service Overview

With the business expansion, many enterprises establish branches in different areas and employees are often on business trips. Therefore, some applications (such as the VoIP, instant messages and network conference) are used widely in enterprises. These applications require a network that supports point-to-multipoint (P2MP) services. In addition, the network reliability must be ensured and a transparent and secure data channel is required for multi-point transmission because of the privacy of the enterprise business data. The VPLS technology is suitable to be deployed in this scenario.

Example Network

Figure 11-9 shows the example network of VPLS enterprise access.

Figure 11-9 Example network of VPLS enterprise access

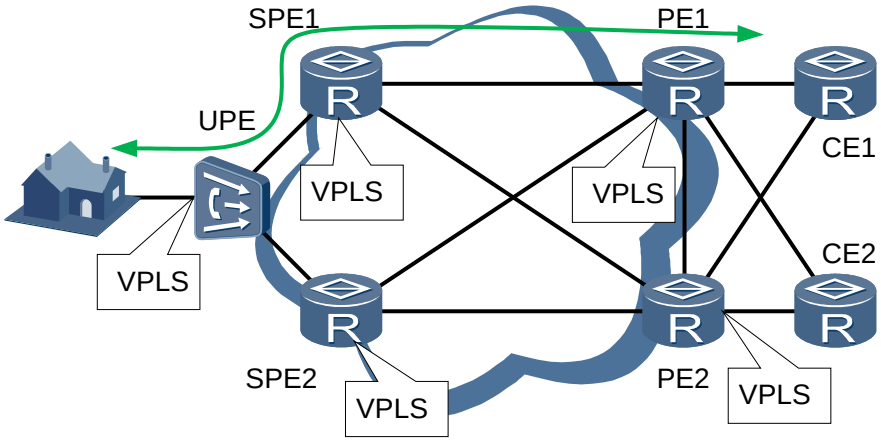


- The virtual private network (VPN) between different branches is achieved by deploying the VPLS.
- The pseudo wire (PW) redundancy is used to protect the important branches (such as branch C in the figure).
- An OLT/MSAN, functioning as the main node, implements the Layer 2 label switching, and other branches are connected to the VPLS network through backup PWs.
- The split horizon between the VPLS and PW is canceled.
- The basic Layer 2 forwarding mechanism in this scenario is consistent with that in the VPLS individual access scenario except that the split horizon needs to be canceled and the PW protection needs to be supported for Layer 2 forwarding in this scenario.

11.4.3 VPLS PW Redundancy for Protecting Multicast Services

Figure 11-10 illustrates an application of VPLS PW redundancy for protecting multicast services, such as Internet Protocol television (IPTV) services, on a hierarchical virtual private LAN service (HVPLS) network.

Figure 11-10 VPLS PW redundancy for protecting multicast services



Multicast sources CE1 and CE2 are each dual-homed to provider edge 1 (PE1) and PE2 through enhanced trunks (E-Trunks); common PWs connect PEs and superstratum PEs (SPEs). A gateway user-end provider edge (UPE) connects the user end to SPE1/SPE2. The link between the UPE and SPE1 and the link between the UPE and SPE2 back up each other.

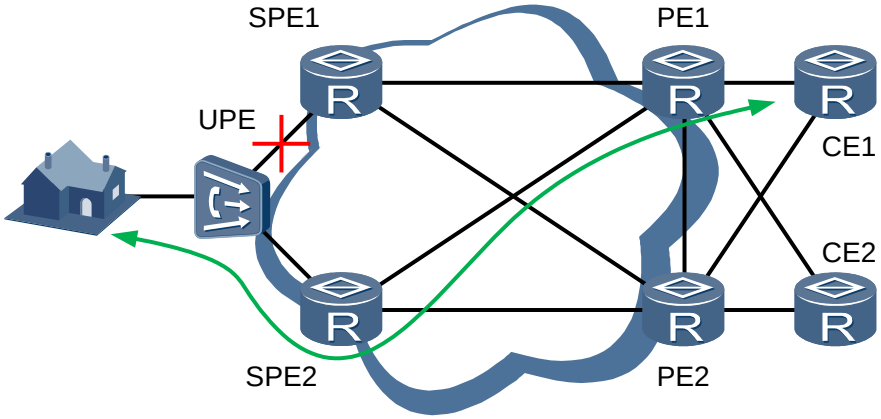
For this networking, the UPE must use PW redundancy in Master/Slave mode because SPE1 and SPE2 do not exchange signaling for determining which one is the master SPE. Upon detecting that the primary PW fails, the UPE rapidly switches traffic to the secondary PW, instructs SPE2 to work as the primary SPE, and sends MAC Withdraw messages to SPE2 instructing SPE2 to delete the MAC addresses learned from SPE1. SPE2 transmits the MAC Withdraw messages to PE1 and PE2, instructing PE1 and PE2 to clear the MAC addresses learned from SPE1. After deleting the MAC addresses learned from SPE1, PE1 will relearn MAC addresses through multicast packets upon receiving traffic from CE1 and CE2 and switch received traffic to the secondary link.

Figure 11-10 shows service traffic when no fault occurs. The following describes how VPLS PW redundancy protects traffic after faults occur.

Primary PW Failure Between the UPE and SPE1

Figure 11-11 shows how traffic is switched if the primary PW between the UPE and SPE1 fails.

Figure 11-11 VPLS PW redundancy protecting services against a failure in the primary PW between the UPE and SPE1



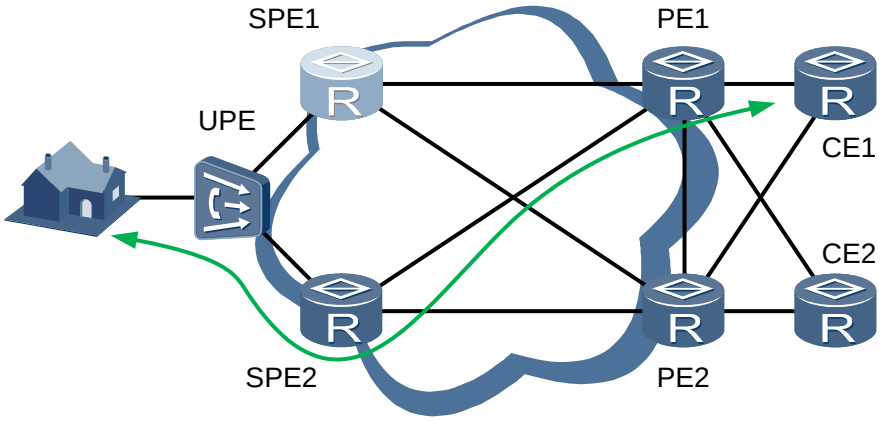
Label switched path (LSP) down events or BFD for PW may cause a PW failure. Upon detecting that the primary PW fails, the UPE switches traffic to the secondary PW and sends MAC Withdraw messages in which the PE ID field carries the SPE1 LSR ID to SPE2. SPE2 transparently transmits the MAC Withdraw messages to PE1 and PE2. SPE2, PE1, and PE2 clear the MAC addresses learned from SPE1.

Switchback: After the primary PW recovers, the UPE instructs SPE2 to change its PW status to secondary and SPE1 to change its PW status to primary. The UPE sends MAC Withdraw messages in which the PE ID field carries the SPE2 LSR ID to SPE1. SPE1 transparently transmits the MAC Withdraw messages to PE1 and PE2. SPE1, PE1, and PE2 clear the MAC addresses learned from SPE2. PE1 and PE2 then relearn MAC addresses through multicast packets from the primary PW.

SPE1 Failure

Figure 11-12 shows how traffic is switched if SPE1 fails.

Figure 11-12 VPLS PW redundancy protecting services against an SPE1 failure



After detecting that SPE1 fails, the UPE switches traffic to the secondary PW and sends MAC Withdraw messages in which the PE ID field carries the SPE1 LSR ID to SPE2. SPE2 transparently transmits the MAC Withdraw messages to PE1 and PE2. SPE2, PE1, and PE2

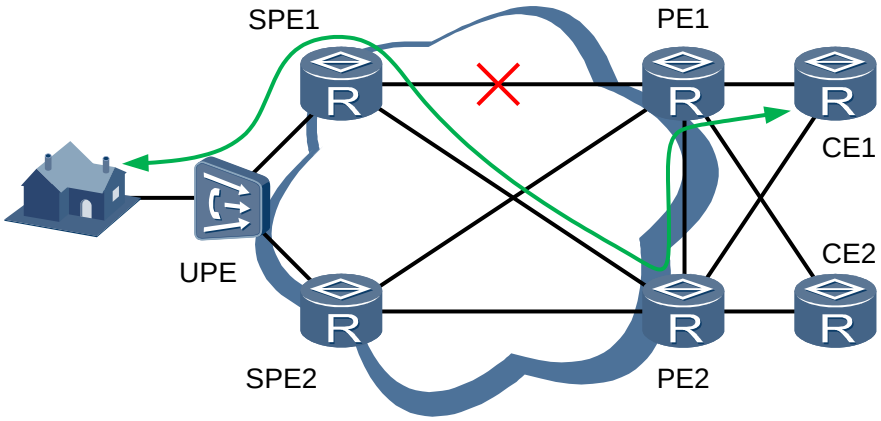
clear the MAC addresses learned from SPE1. Sometimes, PE1 and PE2 detect that the PW passing through SPE1 is faulty before receiving the MAC Withdraw messages and directly clear the MAC addresses learned from SPE1.

Switchback: After the primary PW recovers, the UPE instructs the PW passing through SPE2 to work as a secondary PW and the PW passing through SPE1 to work as the primary PW. The UPE sends MAC Withdraw messages in which the PE ID field carries the SPE2 LSR ID to SPE1. SPE1 transparently transmits the MAC Withdraw messages to PE1 and PE2. SPE1, PE1, and PE2 clear the MAC addresses learned from SPE2. PE1 and PE2 then relearn MAC addresses through multicast packets from the primary PW.

Link Failure Between SPE1 and PE1

Figure 11-13 shows how traffic is switched if the link between SPE1 and PE1 fails.

Figure 11-13 VPLS PW redundancy protecting services against a link failure between SPE1 and PE1



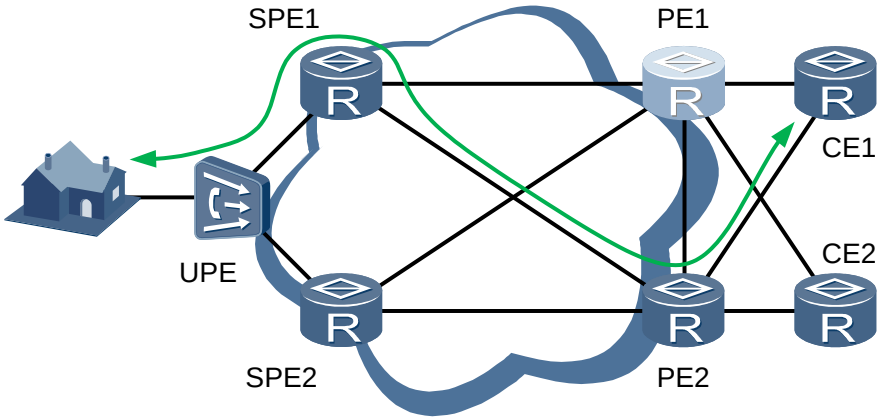
If deployed between SPE1 and PE1, Label Distribution Protocol (LDP) fast reroute (FRR) ensures the availability of traffic between SPE1 and PE1. If LDP FRR is not deployed, LDP LSP ensures the availability of traffic between SPE1 and PE1.

Switchback: Traffic will not be switched back to the primary PW. After LDP LSP convergence, the primary PW is carried by a new LSP.

PE1 Failure

Figure 11-14 shows how traffic is switched if PE1 fails.

Figure 11-14 VPLS PW redundancy protecting services against a PE1 failure



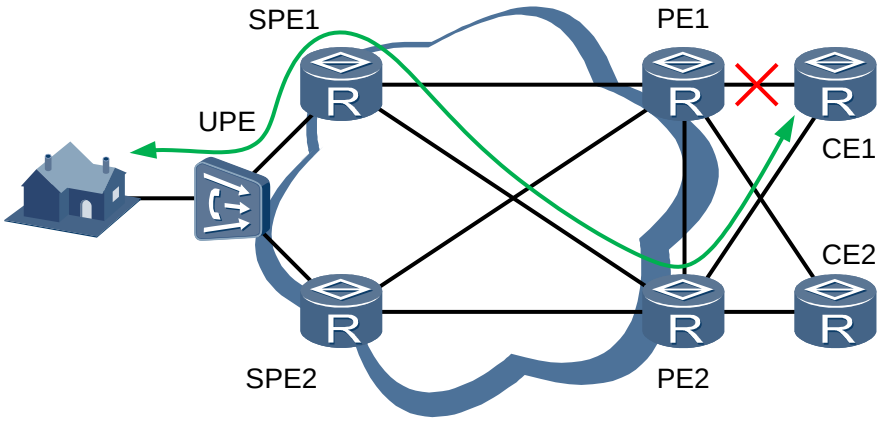
CE1 and CE2 are each dual-homed to PE1 and PE2 through E-Trunk. If PE1 fails, the E-Trunk primary and secondary status changes. PE2 detects E-Trunk status changes and sends MAC Withdraw messages to SPE1 and SPE2, instructing SPE1 and SPE2 to clear MAC addresses in the associated VSI. Sometimes, SPE1 and SPE2 detect that the PW passing through PE1 is faulty before receiving the MAC Withdraw messages and directly clear MAC addresses associated with the PW.

Switchback: If PE1 recovers, traffic switches back to PE1 after a default E-Trunk switchback delay. Upon detecting E-Trunk changes, PE1 and PE2 send MAC Withdraw messages to SPE1 and SPE2, instructing SPE1 and SPE2 to clear MAC addresses learned from PE2.

Primary AC Link Failure

Figure 11-15 shows how traffic is switched if the link between CE1 and PE1 fails.

Figure 11-15 VPLS PW redundancy protecting services against a link failure between CE1 and PE1



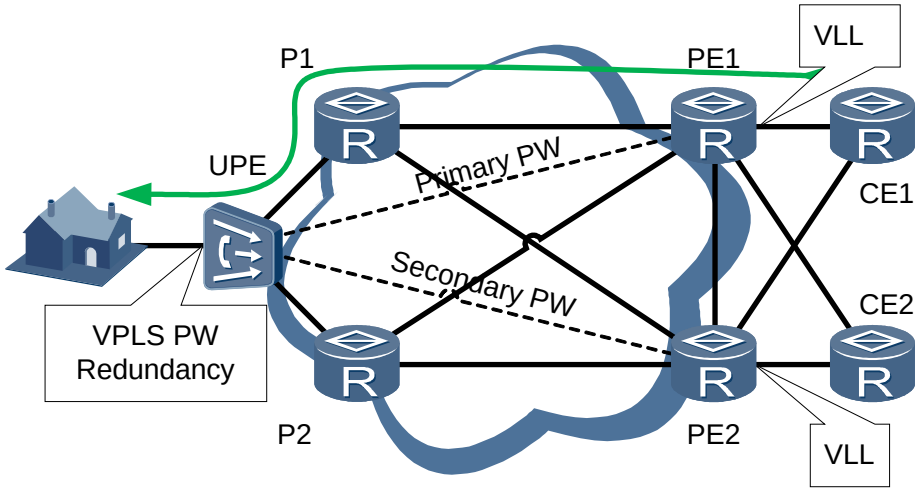
CE1 and CE2 are each dual-homed to PE1 and PE2 through E-Trunk. After the link between CE1 and PE1 fails, the E-Trunk primary and secondary status changes. Upon detecting E-Trunk status changes, PE1 and PE2 send MAC Withdraw messages to SPE1, instructing SPE1 to clear all MAC addresses in the VSI.

Switchback: After the link between CE1 and PE1 recovers, the E-Trunk primary and secondary status changes back. Upon detecting E-Trunk changes, PE1 and PE2 send MAC Withdraw messages to SPE1, instructing SPE1 to clear all MAC addresses.

11.4.4 VPLS PW Redundancy for Protecting Unicast Services

Figure 11-16 illustrates an application of VPLS PW redundancy for protecting unicast services, such as high-speed internet (HSI) or voice over IP (VoIP) services, on a virtual private LAN service (VPLS) and virtual leased line (VLL) interconnected network.

Figure 11-16 VPLS PW redundancy for protecting unicast services



Authentication servers CE1 and CE2 are each dual-homed to PE1 and PE2 through enhanced trunks (E-Trunks). A UPE connects the user end and PE1/PE2. The link between the UPE and PE1 and the link between the UPE and PE2 back up each other.

For this networking, PE1 and PE2 can determine their master and backup statuses through E-Trunk negotiation. Therefore, the UPE can use PW redundancy in Independent mode to determine the PW primary and secondary statuses based on the status of PE1 and PE2. Upon detecting that the primary PW fails, the UPE rapidly switches traffic to the secondary PW and instructs PE2 to work as the master PE. After the E-Trunk detects that the primary PW fails, it switches traffic to the secondary AC link.

Figure 11-16 shows service traffic when no fault occurs. The following describes how VPLS PW redundancy protects services after a fault occurs.

NOTE

When a fault occurs, VPLS PW redundancy protects services differently in Master/Slave and Independent modes. The following describes the difference in service protection provided by VPLS PW redundancy in Master/Slave and Independent modes.

Network Side Failure

Label switched path (LSP) or traffic engineering (TE) tunnels and PW redundancy can protect services against faults on the UPE network side.

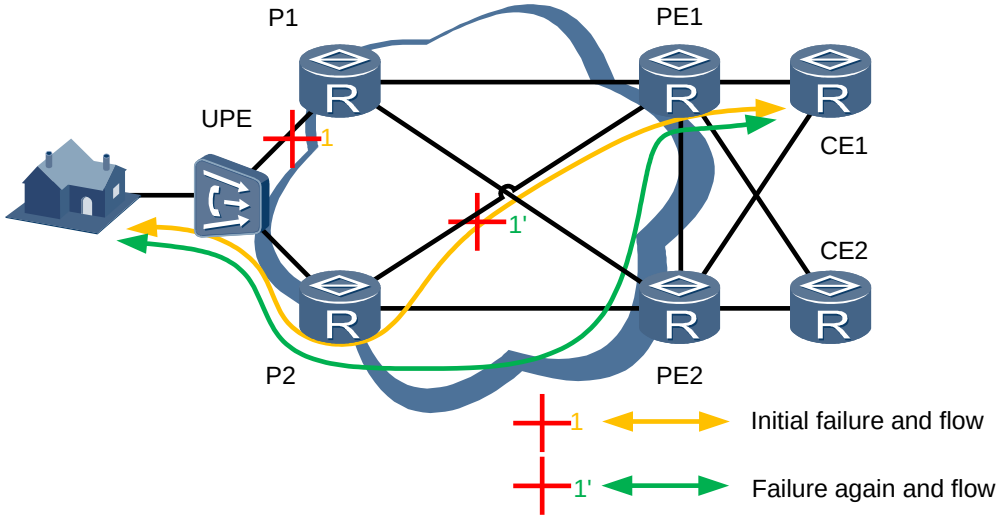
NOTE

If a network-side fault occurs, LSP or TE tunnels detect the fault and switch traffic to other tunnels. If tunnel protection is unavailable or fails, PW redundancy is required to protect traffic. A bypass PW needs to be configured between PE1 and PE2 for PW redundancy.

- LSP or TE tunnel protection: After a network-side fault occurs, routes and LSPs are recomputed. The LSP where the primary PW is located is converged to a new path. Figure 11-17 shows how traffic is switched.

  After the fault is rectified, LSPs re-converge and the primary PW is carried by a new LSP.

Figure 11-17 Protecting services against an LSP failure (bypass PW not configured)



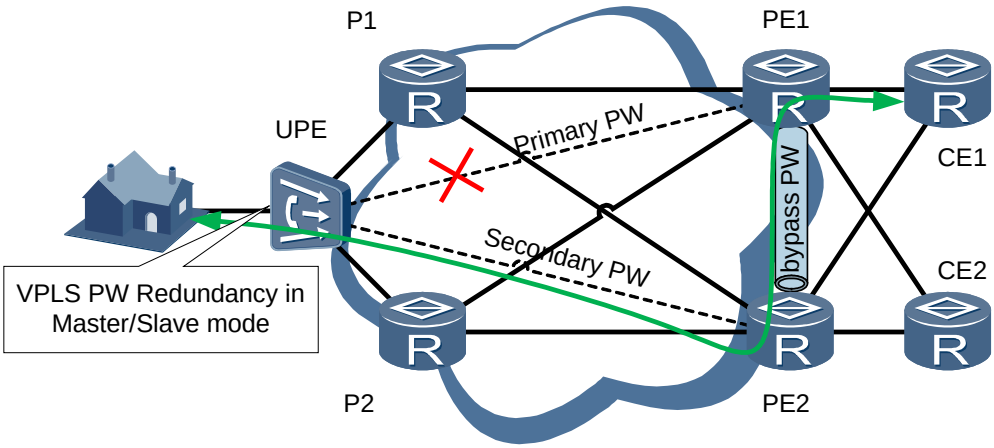
- PW redundancy: If LSP or TE tunnel switching fails, traffic is switched to the secondary PW. Figure 11-18 shows how traffic is switched.

  After the fault is rectified, traffic will be switched back based on preset switchback policies.

NOTE

Bypass PWs are required for PW redundancy to transmit traffic between PE1 and PE2.

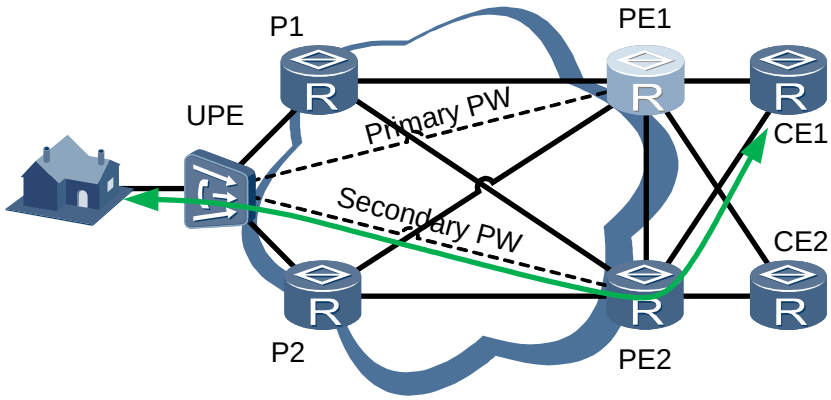
Figure 11-18 Protecting services against a primary PW failure (bypass PW configured)



PE1 Failure

Figure 11-19 shows how traffic is switched if PE1 fails.

Figure 11-19 VPLS PW redundancy protecting services against a PE1 failure



PE2 becomes the master and PE1 becomes the backup after E-Trunk negotiation. The UPE is informed of the switchover. Upon detecting that the primary PW fails, the UPE clears MAC addresses learned from the primary PW and switches traffic to the secondary PW.

Switchback: After the PE1 fault is rectified, PE1 becomes the master through E-Trunk negotiation. Upon detecting PE1 and PE2 status changes, the UPE clears MAC addresses learned from PE2 and relearns MAC addresses through multicast packets.

Primary AC Link Failure

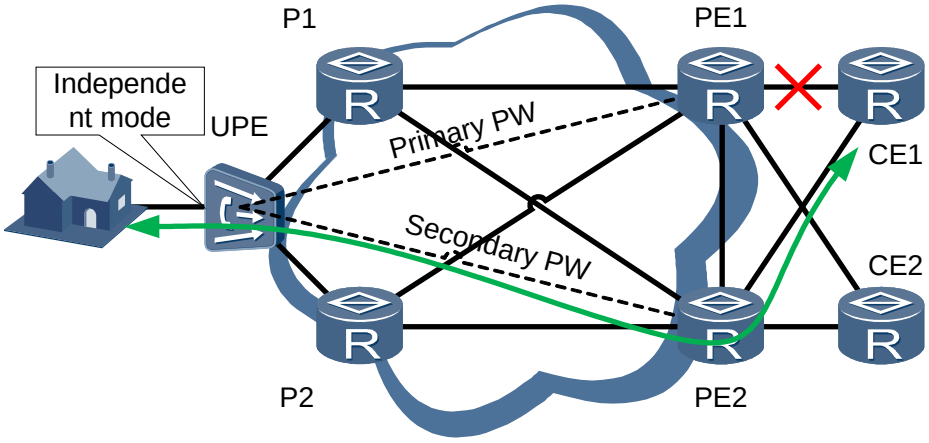
Figure 11-20 shows how traffic is switched if the link between PE1 and CE1 fails.

- In Independent mode

  After the primary AC link between CE1 and PE1 fails, PE2 works as the master after E-Trunk negotiation. The UPE is informed of the switchover. The UPE detects that the primary AC link fails and switches traffic to the secondary PW.

After the link between CE1 and PE1 recovers, PE1 becomes the master after E-Trunk negotiation. Upon detecting PE1 and PE2 status changes, the UPE clears MAC addresses learned from PE2 and relearns MAC addresses through multicast packets.

Figure 11-20 VPLS PW redundancy in Independent mode protecting services against a failure in the primary AC link

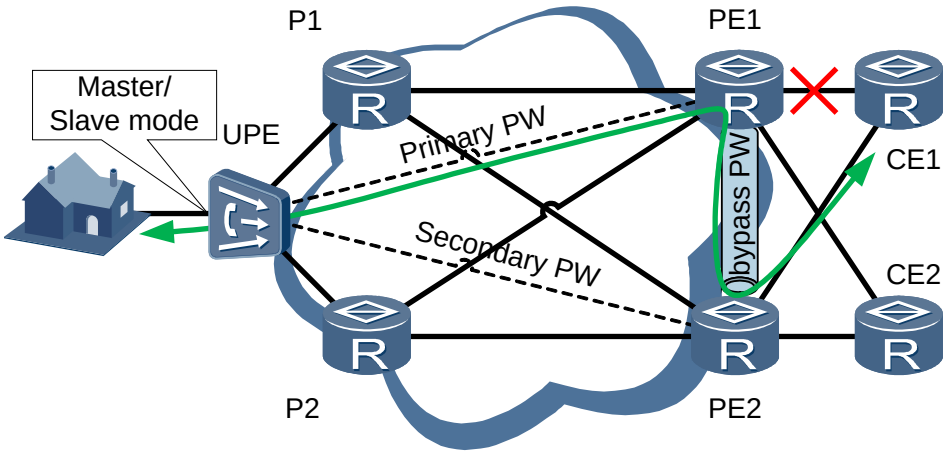


- In Master/Slave mode

  After the primary AC link between CE1 and PE1 fails, PE2 works as the master after E-Trunk negotiation. The PW status on the UPE remains unchanged.

  After the link between CE1 and PE1 recovers, PE1 becomes the master after E-Trunk negotiation. Upon detecting PE1 and PE2 status changes, the UPE clears MAC addresses learned from PE2 and relearns MAC addresses through multicast packets.

Figure 11-21 VPLS PW redundancy in Master/Slave mode protecting services against a failure in the primary AC link



11.5 Configuring VPLS MP2MP Intercommunication

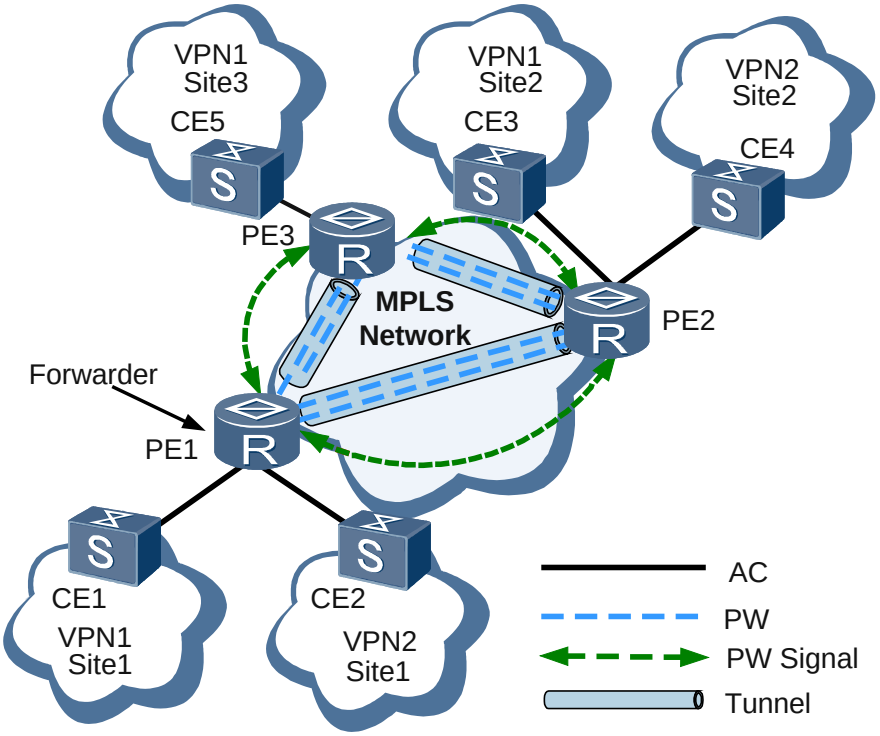
VPLS can implement the multipoint-to-multipoint (MP2MP) VPN networking; therefore, by using the VPLS technology, service providers (SPs) can provide the Ethernet-based multipoint services through MPLS backbone networks.

Application Context

A lot of private line services in carriers' network use the virtual private network (VPN) virtual private wire service (VPWS) technology, which can provide point-to-point (P2P) communication services on Layer 2 or Layer 3 network. With the development of Ethernet and MPLS technologies, carriers hope to provide not only P2P services on the private network, but Ethernet-like point-to-multipoint (P2MP) services on the metropolitan area network (MAN) and wide area network (WAN). By deploying virtual private LAN service (VPLS) technology on the provider edge (PE), carriers can provide Ethernet-based MP2MP services for users through MPLS backbone networks, achieving the local area network (LAN) simulation.

Figure 11-22 shows the basic VPLS transmission process. Full-meshed PWs are created through signaling transmission by PE routers. Transmission of packets between CEs relies on VSIs configured on PEs, and PWs established between the VSIs.

Figure 11-22 Basic VPLS transmission process



Prerequisite

1. The IP address of the loopback interface must be configured.
2. The LSR ID must be configured.
3. The VLAN for MPLS label forwarding must be created.

4. The global MPLS, VLAN MPLS, and VLAN interface MPLS must be enabled.
5. MPLS L2VPN must be enabled.
6. A static or dynamic route must be successfully configured on each device in the network (so that LSRs can reach each other through the IP route).
7. The global LDP function is enabled and remote LDP sessions are configured.
8. The vlan-based traffic stream must be created.

Data preparation

Before configuring the VPLS P2MP intercommunication services, plan the data items as listed in Table 11-7.

Table 11-7 Plan of VPLS P2MP intercommunication service data items

Item	Data	Remarks
MA5600T/MA5603T/MA5608T	VSI	VLANs are mapped to the VPLS domain after a VSI is bound to the VLAN and PW. VLAN mapping allows service packets to be broadcast in the VPLS domain.
	VPLS PW	-

Procedure

Configure a VSI.

VSIs are the core of VPLS services. With VSIs, actual links carrying VPLS services can be mapped into PWs.

1. In global config mode, run the **vsi** command to create a VSI and enter the VSI mode.
2. Run the **pwsignal ldp** command to configure the signaling type for VSI as LDP.

   Currently, you can only configure the signaling type for VSI as LDP.
3. Run the **vsi-id** command to configure the VSI ID.

   Once the VSI ID is successfully set, it cannot be changed or deleted. If you need to change it, delete the VSI.

Step 1 (Optional) Configure VSI attributes.

In VSI mode, configure VSI basic attributes based on actual requirements. VSI basic attributes include the VSI description information, encapsulation type, control words, maximum transmission unit (MTU), and traffic suppression.

- Run the **description** command to configure the description of a VSI.
- Run the **encapsulation** command to configure the encapsulation type of a VSI.
- Run the **control-word** command to enable the control word of a VSI. After the control word is enabled, control information will be added to packets.

NOTE

If you use the **control-word** command in VSI mode and the **control-word** command in PW-para-index mode to configure the control word concurrently, the one set by the **control-word** command in PW-para-index mode takes effect.

- Run the **mtu** command to set the MTU of a VSI.
- Run the **traffic-suppress** command to set the suppression level of the broadcast, unknown multicast, and unknown unicast traffic for a VSI.

  Before configuring the multicast service carried in VPLS, you must disable the VSI unknown multicast supression. Otherwise, packet loss will occur in the multicast services.

Step 2 Configure PWs.

1. In global config mode, run the **pw-para pwindex** *pwindex* command to create a PW and enter the PW-para-index mode.

   For a VPLS PW, you must first create the PW-para-index mode and then perform the PW binding.
2. Run the **service-type vpls** command to configure the service type of a PW as VPLS.
3. Run the **pwid** command to configure the ID of a PW.
4. Run the **peer-address** command to set the IP address of the peer device of a PW.
5. Run the **pw-type ethernet** command to configure the type of a PW as Ethernet.

   When the service type is VPLS, you can set the PW type only to **Ethernet**. The PW type must be identical to the VSI encapsulation type.
6. (Optional) Run the **control-word** command to enable the control word of a PW.
7. Run the **dyn-receive-label** command to specify the incoming label of a dynamic PW.

Step 3 In VSI mode, run the **vsi-pw-binding** command to bind the VSI to the PW to create a VPLS PW service.

Step 4 In VSI mode, run the **vsi-ac-binding vlan** command to bind a VLAN to the VSI.

After the above configurations are complete, VLAN service packets can be forwarded within a VSI.

----End

Example

Assume that VLAN 100 is used for MPLS forwarding, a VSI and a PW are created, and the PW and VLAN 100 are bound to the VSI respectively. To configure VSI and PW parameters as follows:

- To set the VSI ID to **1**, the VSI name to **hsi**, and the signaling mode to **LDP**, and retain the default values for other parameters, do as follows:
- To set the PW index to **1**, the service type to **VPLS**, the PW ID to **1**, the IP address of the peer device to **1.1.1.1**, the encapsulation type to **Ethernet tagged**, and the dynamic PW incoming label to **10240**, do as follows:

```
huawei(config)#vsi hsi
huawei(config-vsi-hsi)#pwsignal ldp
huawei(config-vsi-hsi)#vsi-id 1
huawei(config-vsi-hsi)#quit
huawei(config)#pw-para pwindex 1
huawei(config-pw-para-index-1)#service-type vpls
```

```
huawei(config-pw-para-index-1)#pwid 1
huawei(config-pw-para-index-1)#peer-address 1.1.1.1
huawei(config-pw-para-index-1)#pw-type ethernet tagged
huawei(config-pw-para-index-1)#dyn-receive-label 10240
huawei(config-pw-para-index-1)#quit
huawei(config)#vsi hsi
huawei(config-vsi-hsi)#vsi-pw-binding pwindex 1
huawei(config-vsi-hsi)#vsi-ac-binding vlan 100
```

12 Layer 2 VPN

About This Chapter

Layer 2 VPN technology refers to PW technologies and TDM emulation transmission technologies implemented by the system.

12.1 PWE3

PWE3 is an end-to-end Layer 2 service carrying technology and is a type of point-to-point L2VPN technology. PWE3 is proposed by the IETF PWE3 working group as one of the solutions for connecting the traditional communication network with the PSN network.

12.1.1 Introduction

Definition

Pseudo-wire emulation edge to edge (PWE3) is a type of Layer 2 service carrying technology. It is mainly used to emulate the essential behavior and characteristics of the services such as the ATM, frame relay, Ethernet, low-rate time division multiplexing (TDM) circuit, and synchronous optical network (SONET)/synchronous digital hierarchy (SDH) as faithfully as possible in a packet switched network (PSN).

PWE3 is implemented on access devices through MPLS and IP technologies. MPLS supports PWE3 by using the LDP or RSVP-TE protocol as signaling.

Purpose

PWE3 can interconnect the traditional network with PSN to share resources and expand the reach of networks. For example, PWE3 can emulate services such as TDM, ATM, and Ethernet, and can implement service interoperation by using existing PSN (IP/MPLS) as the bearer network.

Benefit

PWE3 connects the traditional TDM, ATM, and Ethernet networks with PSN (IP/MPLS). In this way, PWE3 protects the investment on the traditional TDM, ATM, and Ethernet networks, and also implements the all-IP network architecture.

12.1.2 Reference Standards and Protocols

The following lists the reference standards and protocols of this feature.

- RFC 3985: Pseudo Wire Emulation Edge-to-Edge (PWE3) Architecture
- RFC 4447: Pseudowire Setup and Maintenance Using the Label Distribution Protocol (LDP)
- RFC 3916: Requirements for Pseudo-Wire Emulation Edge-to-Edge (PWE3)
- RFC 4446: IANA Allocations for Pseudo wire Edge to Edge Emulation (PWE3)
- RFC 4717, Encapsulation Methods for Transport of Asynchronous Transfer Mode (ATM) over MPLS Networks
- RFC 4816, Pseudowire Emulation Edge-to-Edge (PWE3) Asynchronous Transfer Mode (ATM) Transparent Cell Transport Service
- RFC 4448: Encapsulation Methods for Transport of Ethernet over MPLS Networks
- RFC 5085: PW vccv A control Channel for PW
- RFC 4553: Structure-Agnostic Time Division Multiplexing (TDM) over Packet (SAToP)
- RFC 5462: Multi-Protocol Label Switching (MPLS) Label Stack Entry EXP Field Renamed to Traffic Class Field
- RFC 4385: Pseudowire Emulation Edge-to-Edge (PWE3) Control Word for Use over an MPLS PSN
- draft-ietf-pwe3-redundancy-bit-00.txt
- draft-bryant-filsfils-fat-pw-03.txt

12.1.3 Principle

Basic Principle of PWE3

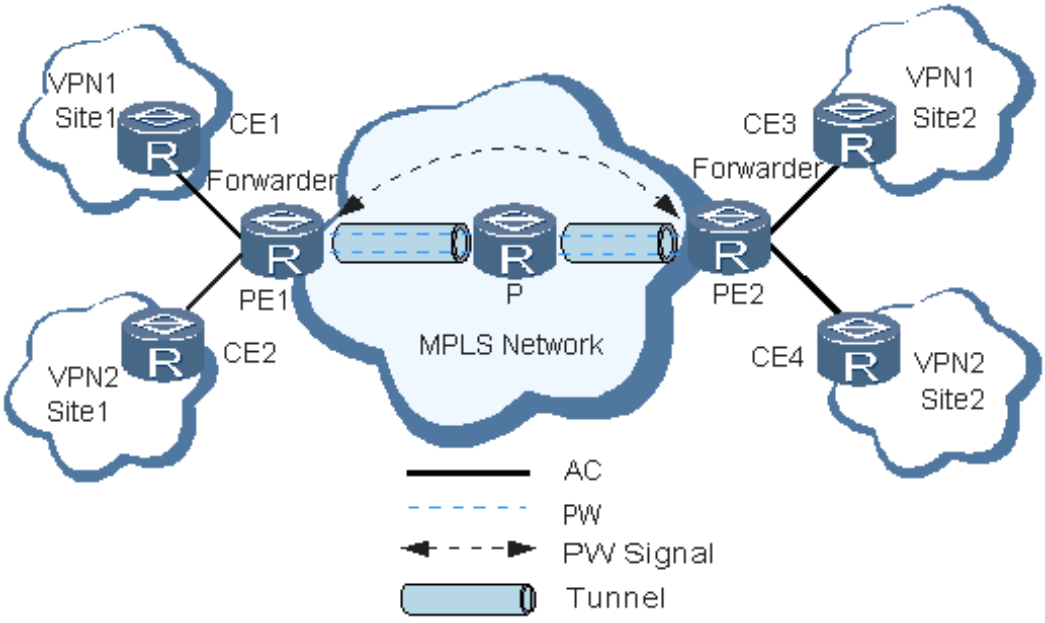
Basic PWE3 Transmission Components

Pseudo wire emulation edge-to-edge (PWE3), which uses LDP and RSVP-TE as the signaling protocols, carries various types of Layer 2 services, such as various types of Layer 2 data packets, from the customer edge (CE), and transparently transmits the Layer 2 data through tunnels (such as MPLS LSP or TE tunnels). As shown in Figure 12-1, the basic PWE3 transmission components include the following:

- Attachment circuit (AC): a link between CE and PE. All user packets (including Layer 2 and Layer 3 protocol packets of users) on the AC are transparently forwarded to the peer end.
- Pseudo wire (PW): PW is a kind of virtual connection between two PEs. It is a mechanism for two PEs in the packet-switched network to transfer essential elements of emulation service. A PW conveys VC information by signaling (LDP or RSVP-TE). Since VC is directional, PW is directional too. For the PWE3 system, a PW is like a direct channel between a local AC and a peer AC and is used for transparently transmitting the Layer 2 data of users.
- Forwarder: After a PE receives data frames from an AC, the forwarder selects a PW for forwarding the frames. In fact, the forwarder is a forwarding table of PWE3.
- Tunnel: A tunnel is a direct channel between a local PE and a peer PE and is used for transparently transmitting data between the PEs. Tunnels are used for carrying PWs. A tunnel can carry multiple PWs. Generally, the tunnel refers to an MPLS tunnel.

- PW signaling protocol: A PW signaling protocol is the basis for implementing PWE3 and is used to create and maintain PWs. Current PW signaling protocols are mainly LDP and RSVP-TE.
- Encapsulation: The packets transmitted through the PW use the standard PW encapsulation format and technology. There are multiple PWE3 encapsulation types on a PW. The formats are defined in detail in draft-ietf-pwe3-iana-allocation-x.
- Quality of service (QoS): The priority information at the header of Layer 2 user packets is mapped to the QoS priority for transmitting the packets in the public network. In general, support for MPLS QoS is required.

Figure 12-1 Basic PWE3 transmission components



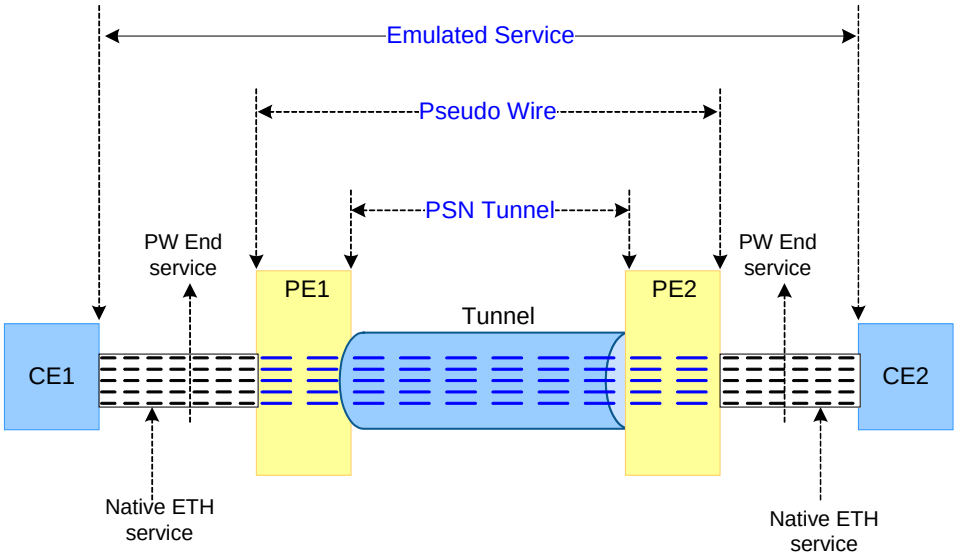
Assume that the VPN1 packet stream travels from CE1 to CE3. The basic data flow would be as follows:

- CE1 transmits a Layer 2 packet to PE1 through an AC.
- After PE1 receives the packet, the forwarder selects a PW for forwarding the packet.
- PE1 generates two MPLS labels according to the PW forwarding entry. The private network label is used for identifying the PW, and the public network label is used for transmitting the packet to PE2 through the tunnel.
- The Layer 2 packet arrives at PE2 through the public network tunnel. The system extracts the private network label (the public network label is extracted by the penultimate P device).
- The forwarder of PE2 selects an AC for forwarding the packet, and then PE2 forwards the packet to CE3.

PWE3 Network Model

Figure 12-2 shows a PWE3 reference model.

Figure 12-2 PWE3 network model



The channel set up in a PWE3 network is a point-to-point channel. Channels are isolated from each other. Layer 2 user packets are transparently transmitted between PWs. The following provides a detailed description.

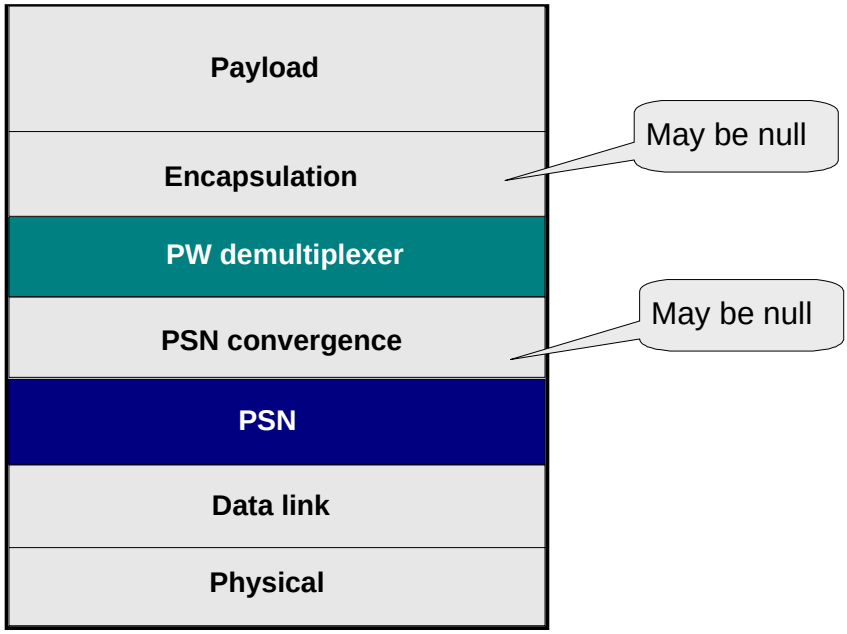
- According to the services requirements of the CE, one or more PWs are set up between PE1 and PE2. Multiple PWs can be carried on one PSN tunnel.
- For the PEs, after the PW is set up, the mapping between the user access interface (AC) and virtual link (PW) is determined.
- The PSN device only needs to forward the MPLS packet according to the MPLS label, regardless of the Layer 2 user packet encapsulated inside the MPLS packet.

PWE3 Service Model

Figure 12-3 shows a PWE3 service model. According to the PWE3 service model, PWE3 is presented by an outer label (PSN tunnel) and an inner label (PW demultiplexer).

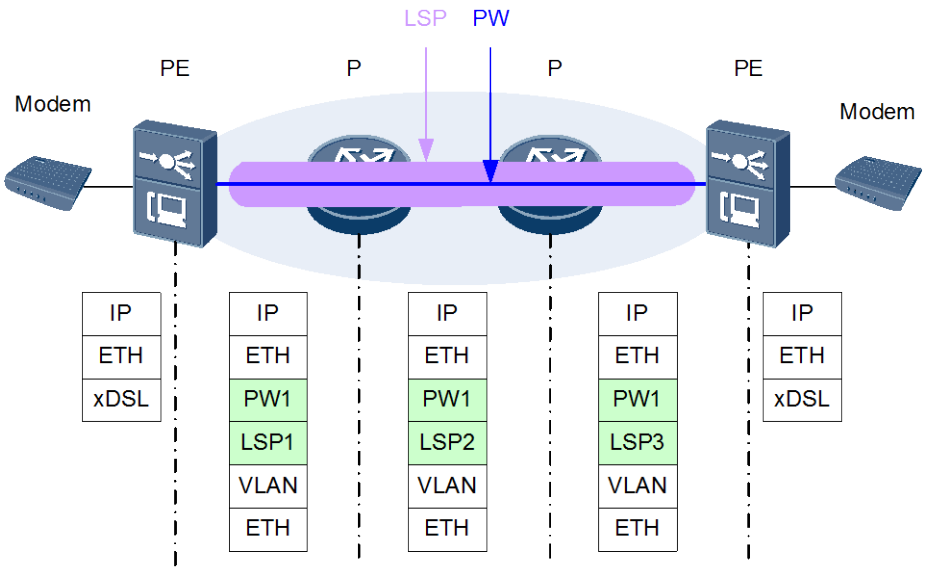
The PSN layer can adopt the MPLS and IP technologies, and the PW demultiplexer layer can adopt the MPLS, UDP, or L2TP technology. Hence, the supported combinations of PWE3 outer labels and inner labels are as follows: MPLS over MPLS, MPLS over IP, UDP over IP, and L2TP over IP. The MA5600T/MA5603T/MA5608T currently supports the first three combinations.

Figure 12-3 PWE3 service model



- Figure 12-4 shows the PWE3 protocol stack in the MPLS over MPLS encapsulation mode.

Figure 12-4 PWE3 protocol stack in the MPLS over MPLS encapsulation mode

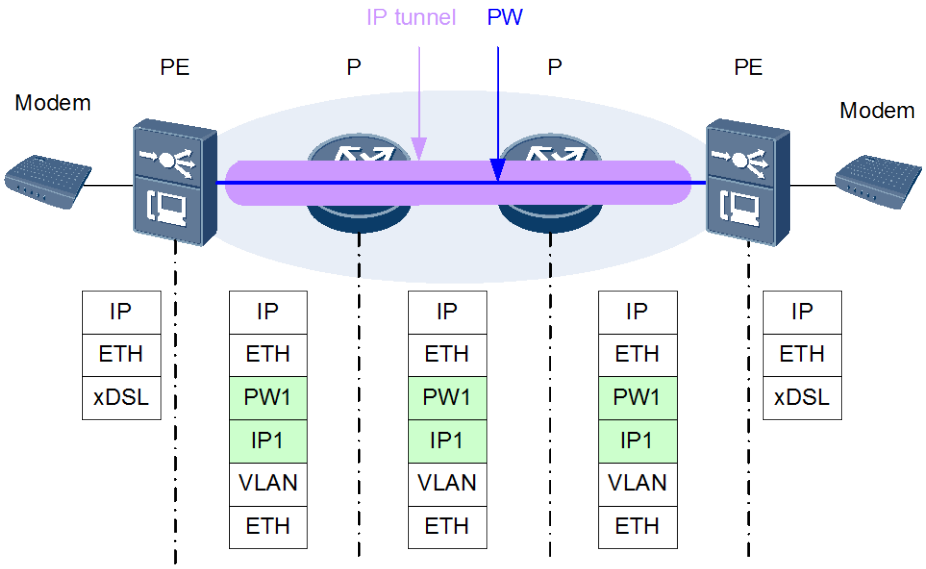


PW: PW label (inner label), identifying user service

LSP: MPLS label (outer label), identifying PSN tunnel

- Figure 12-5 shows the PWE3 protocol stack in the MPLS over IP encapsulation mode.

Figure 12-5 PWE3 protocol stack in the MPLS over IP encapsulation mode

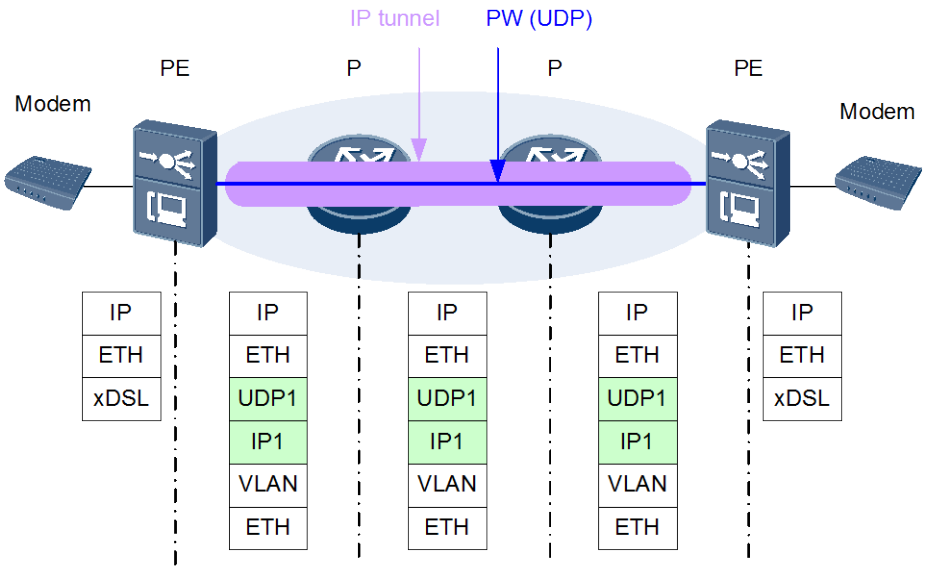


PW: PW label (inner label), identifying user service

IP: IP tunnel (outer label), identifying PSN tunnel

- Figure 12-6 shows the PWE3 protocol stack in the UDP over IP encapsulation mode.

Figure 12-6 PWE3 protocol stack in the UDP over IP encapsulation mode

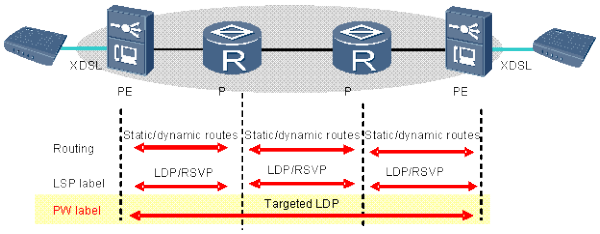


UDP: UDP port number (inner label), identifying user service

IP: IP tunnel (outer label), identifying PSN tunnel

Figure 12-7 illustrates the principle of PW label distribution.

Figure 12-7 Principle of PW label distribution



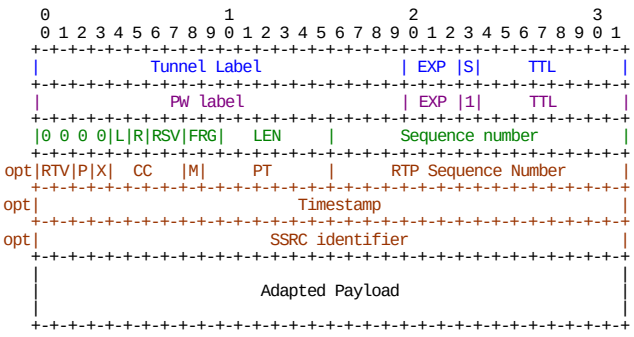
Packet Format

PWE3 has three packet formats: MPLS+PW, IP+PW, and IP+UDP. The MA5600T/MA5603T/MA5608T currently supports MPLS+PW and IP+PW.

- MPLS+PW: In this packet format, the combination of PWE3 outer label and inner label is MPLS over MPLS. It is applicable to MPLS network transmission.

  Figure 12-8 shows the format of an MPLS+PW PWE3 packet.

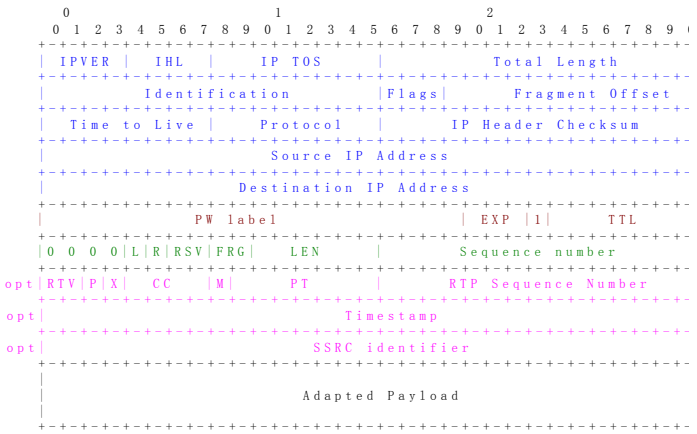
  Figure 12-8 Format of an MPLS+PW PWE3 packet



- IP+PW: In this packet format, the combination of PWE3 outer label and inner label is MPLS over IP. It is applicable to MPLS over IP network transmission. Different from MPLS+PW packets, IP+PW packets are forwarded at IP Layer 3 when transmitted over the PSN network.

  Figure 12-9 shows the format of an IP+PW PWE3 packet.

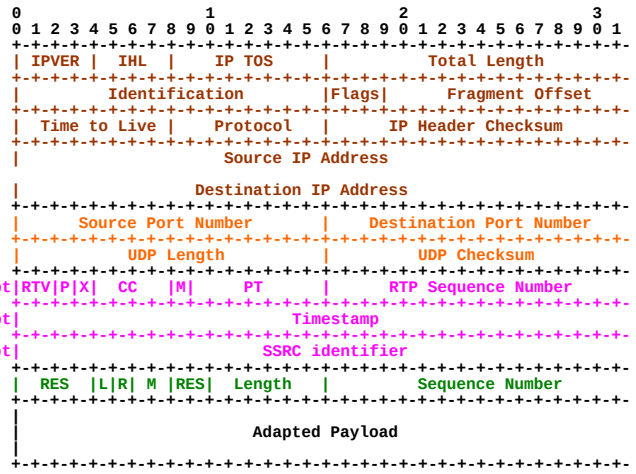
Figure 12-9 Format of an IP+PW PWE3 packet



- IP+UDP: In this packet format, the combination of PWE3 outer label and inner label is UDP over IP. It is applicable to IP network transmission. Different from MPLS+PW packets, IP+UDP packets are forwarded at IP Layer 3 and meanwhile forwarded at Layer 2 by UDP port redirection when transmitted over the PSN network.

  Figure 12-10 shows the format of an IP+UDP PWE3 packet.

Figure 12-10 Format of an IP+UDP PWE3 packet



Principle of TDM PWE3

The TDM service is transmitted over the PSN by circuit emulation. there are two packet encapsulation formats.

- The structure-aware packet, which is also called CESoPSN packet, is defined in RFC5086 and RFC5687.

NOTE

Circuit emulation service over packet switched network (CESoPSN) is a generic term for circuit emulation service and also refers to structure-aware circuit emulation.

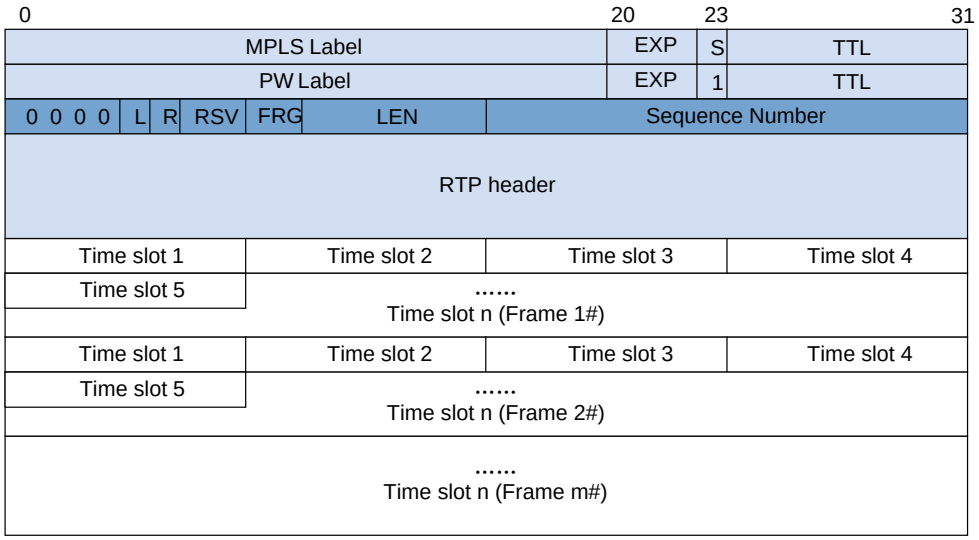
- The structure-agnostic packet, which is also called SAToP packet, is defined in RFC4553.

CESoPSN Packet Format (structure-aware packet format)

The CESoPSN standard provides the channelized TDM service with emulation and transmission functions, and can identify the TDM frame structure and in-frame signaling. Therefore, if the customer needs to provide services based on the timeslot, the CESoPSN packet format can meet this requirement.

Figure 12-11 shows the format of the CESoPSN packet.

Figure 12-11 CESoPSN packet format



- A CESoPSN packet contains a four-byte MPLS header and a four-byte PW header. The length of the CESoPSN control word is 4 bytes, including fields as shown in Table 12-1.
- The length of the Real-Time Transport Protocol (RTP) header is 12 bytes, including the version number, padding flag, and time stamp fields. The time stamp field, whose length is 32 bits, is used for clock synchronization. For format of the RTP header, see RFC3550.
- Time slot indicates the TS in the TDM frame. Each TS occupies 8 bits. All TSs comprise the encapsulated TDM data payload, which does not include the CRC bit. The number of encapsulated frames and the number of TSs in each frame can be set by users according to conditions.

Table 12-1 Fields of the control word

Field	Description
0000	This field is generally all 0s, with the length of 4 bits. When the virtual circuit connectivity verification (VCCV) is needed to help to monitor the SAToP PW status, these four bits are used to identify the start of the associated channel header (ACH).
L	Indicates whether the TDM data in the packet is valid. Its length is

Field	Description
	1 bit. When it is set to 1, it indicates that the TDM data in the packet is invalid; that is, the TDM data in the packet can be neglected to save bandwidth resources.
R	Indicates whether the interconnection function of the local customer edge (CE) is in the packet loss state. When it is set to 0, it indicates that consecutive packets have been received and will no longer be lost.
RSV	Indicates the reserved bit. Its length is 2 bits.
FRG	Indicates the fragmentation status of the packet. Its length is 2 bits. Its values are as follows: • 00: Indicates that the packet encapsulates the entire TDM data. • 01: Indicates that the packet encapsulates the first fragmentation of the TDM data. • 10: Indicates that the packet encapsulates the last fragmentation of the TDM data. • 11: Indicates that the packet encapsulates the intermediate fragmentation of the TDM data.
LEN	Indicates the length of the entire CESoPSN packet (the size of the CESoPSN header and TDM data.) When the length is shorter than 64 bytes, LEN is a specific length value. When the length is equal to or longer than 64 bytes, LEN is 0.
Sequence Number	Indicates the sequence number for transmitting the CESoPSN packet. Its length is 16 bits. Its initial value should be random. It must be incremented by 1 with each CESoPSN data packet sent in the specific PW.

The CESoPSN packet structure has the following characteristics:

- CESoPSN provides emulation and transmission of structure-aware TDM service. That is, CESoPSN can identify the TDM frame structure and in-frame signaling and transmit the frames.

  For example, a structure-aware E1 link consists of 32 timeslots. Except timeslot 0, the other 31 timeslots can each carry a channel of 64 kbit/s voice service. Timeslot 0 is used only for transmitting signaling and frame delimiter.
- CESoP can identify the TDM frame structure. Therefore, idle timeslots are not transmitted, and only the data in the timeslots that are useful for the CE devices is retrieved from the E1 service stream and transmitted after being encapsulated into CESoPSN frames.
- CESoPSN can identify and transmit the CAS and CCS signaling of E1 service streams.

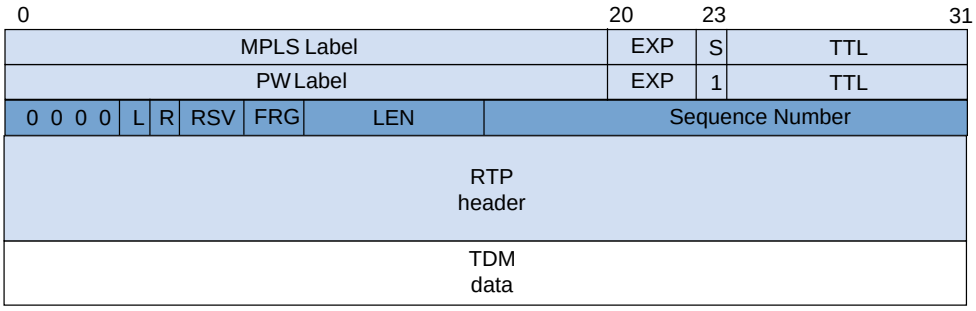
SAToP Packet Format (structure-agnostic packet format)

The structure-agnostic TDM over PSN (SAToP) standard provides the non-channelized TDM service with the emulation and transmission functions. The protocol need not be aware of the structure of the TDM packets and transparently transmits the packets. Therefore, if the

customer only needs to provide services based on E1/T1, SAToP (unstructured packet format) can meet this requirement.

Figure 12-12 shows the format of the SAToP packet.

Figure 12-12 SAToP packet format

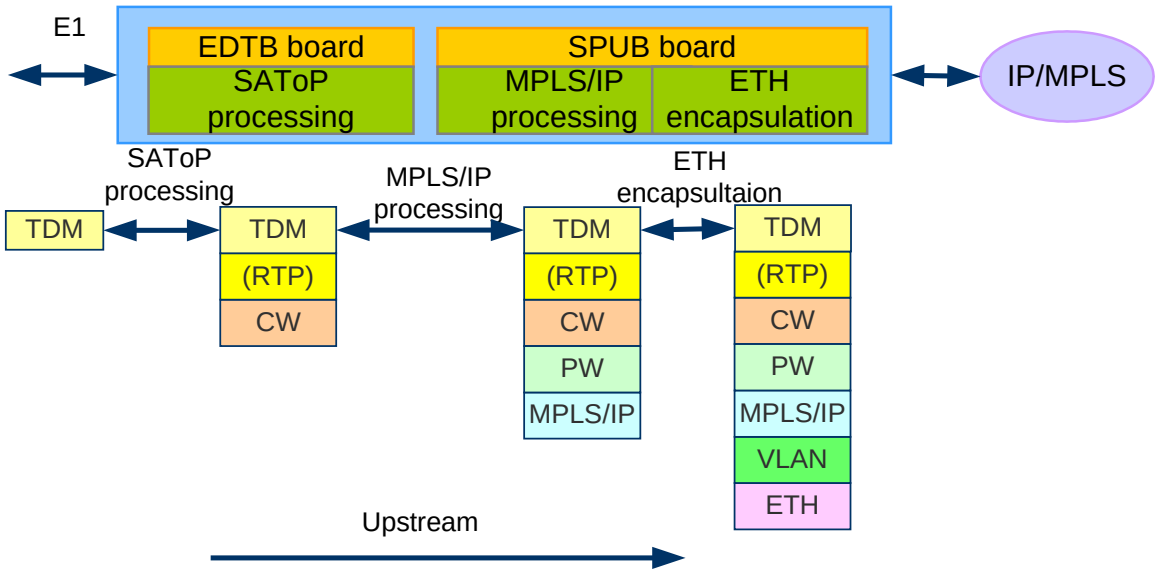


- An SAToP packet also contains a four-byte MPLS header and a four-byte PW header. The length of the SAToP control word is 4 bytes, including fields as shown in Table 12-1.
- The SAToP protocol treats the TDM service as serial data code stream for segmentation, and transmits the service over PWs after encapsulation. SAToP can transmit the synchronization timing information although it is unaware of the structure of the TDM frame.

Service Processing Flow - E1 Access, SAToP Encapsulation, and Ethernet Upstream Transmission

The MA5600T/MA5603T/MA5608T supports E1 access service by EDTB board, and also supports SAToP encapsulation and processing of E1 service. Figure 12-13 shows the service processing flow.

Figure 12-13 E1 access, SAToP encapsulation, and Ethernet upstream transmission



1. Packing/Unpacking of SAToP packets

   The MA5600T/MA5603T/MA5608T packs E1 data in the SAToP format, and adds the control word and RTP header (optional in the MPLS mode) to the SAToP packets.

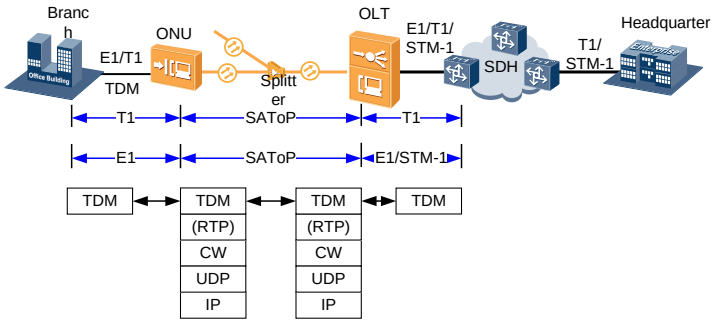
2. Encapsulation of MPLS labels
   - The MA5600T/MA5603T/MA5608T adds/deletes the MPLS labels, and maps inner labels to user circuits.
   - In the MPLS+MPLS encapsulation, the outer LSP label is used for transmitting the packet over an MPLS network; in the IP+MPLS encapsulation, the outer IP address is used for transmitting the packet over an IP network. The inner label is used for mapping to a user circuit.
   - The inner PW tunnel is a bidirectional MPLS tunnel that carries TDM data. A PW label can be statically configured or dynamically created through protocol (LDP).
   - The outer tunnel can be MPLS-encapsulated or IP-encapsulated. In the case of MPLS encapsulation, the outer MPLS tunnel can be statically configured or dynamically created through protocol (LDP or RSVP-TE). In the case of IP encapsulation, the outer IP tunnel can be statically configured.

3. Ethernet processing: In the upstream direction, the ETH header is encapsulated to the packet label header, and then the packet is transmitted through the upstream port on the control board.
   - The upstream VLAN of the TDM PWE3 packet is a service VLAN, which is the VLAN of the corresponding upstream port.
   - The Layer 3 interface MAC address is filled in as the source MAC address of the TDM PWE3 upstream packet, and the MAC address of the next-hop interface (this MAC address can be learned through ARP) is used as the destination MAC address.

Service Processing Flow - Integrated E1/T1 Access in SAToP Mode for the ONU, and SAToP Termination on the OLT of the Same Node

In enterprise private line services, the ONU receives enterprise TDM data through its E1/T1 port, encapsulates the data using SAToP, and sends the SAToP-encapsulated data to the MA5600T/MA5603T/MA5608T through a GPON line. The MA5600T/MA5603T/MA5608T terminates SAToP encapsulation and restores TDM signals. Figure 12-14 shows the service processing flow.

Figure 12-14 Integrated E1/T1 access in SAToP mode for the ONU, and SAToP termination on the OLT of the same node



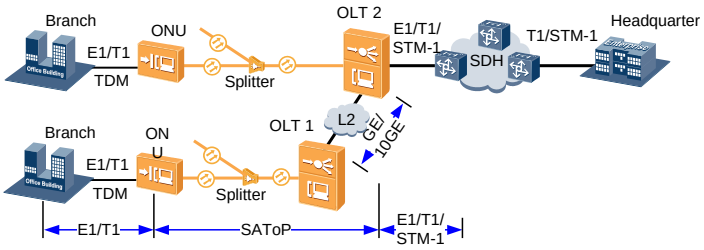
1. The ONU specifies ports and timeslots for E1/T1 services, and encapsulates TDM data in SAToP mode (IP+UDP).

2. The ONU encapsulates the SAToP packets as payload into GEM frames, and sends the GEM frames to the OLT.
3. The OLT decapsulates GEM frames and forwards the extracted SAToP packets to the EDTB board for processing.
4. The EDTB board decapsulates the SAToP packets, restores E1/T1 signals, and sends the data upstream to the SDH network.

Service Processing Flow - Integrated E1/T1 Access in SAToP Mode for the ONU, and SAToP Termination on the OLT of a Different Node

Carrier SDH networks are gradually migrated. In a city, services of only some nodes can be sent upstream through STM-1 ports, and the other nodes need to be connected to regions that have SDH resources using GE/10GE ports. In this case, the ONU provides integrated E1/T1 access in SAToP mode and sends services to the OLT of the same node through a GPON line. This OLT then transparently transmits services to another OLT that has SDH resources using a GE/10GE port. Figure 12-15 shows the service processing flow.

Figure 12-15 Integrated E1/T1 access in SAToP mode for the ONU, and SAToP termination on the OLT of a different node

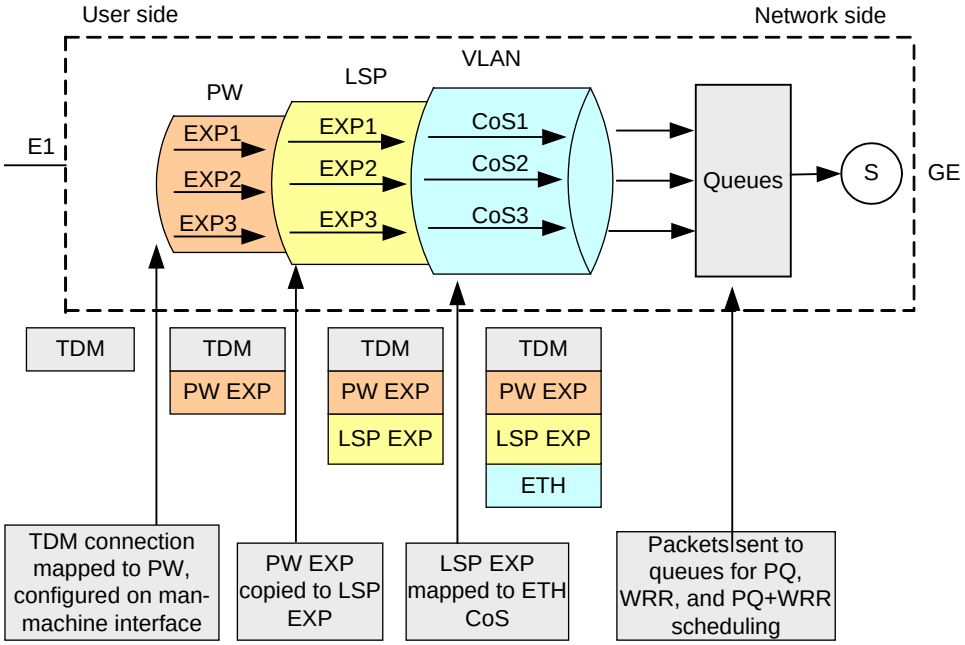


1. The ONU specifies ports and timeslots for E1/T1 services, and encapsulates TDM data in SAToP mode (IP+UDP, with destination IP address and destination UDP port mapped to OLT 2).
2. The ONU encapsulates the SAToP packets as payload into GEM frames, and sends the GEM frames to OLT 1.
3. OLT 1 decapsulates GEM frames, adds Ethernet headers to the extracted SAToP packets, and transparently transmits the packets to OLT 2.
4. OLT 2 restores the SAToP packets and forwards the packets to the EDTB board for processing.
5. The EDTB board decapsulates the SAToP packets, restores E1/T1 signals, and sends the data upstream to the SDH network.

QoS Processing

Figure 12-16 shows the QoS processing flow of TDM PWE3 service, considering the example of SAToP (MPLS over MPLS encapsulation).

Figure 12-16 QoS processing flow of upstream SAToP service (MPLS over MPLS encapsulation)

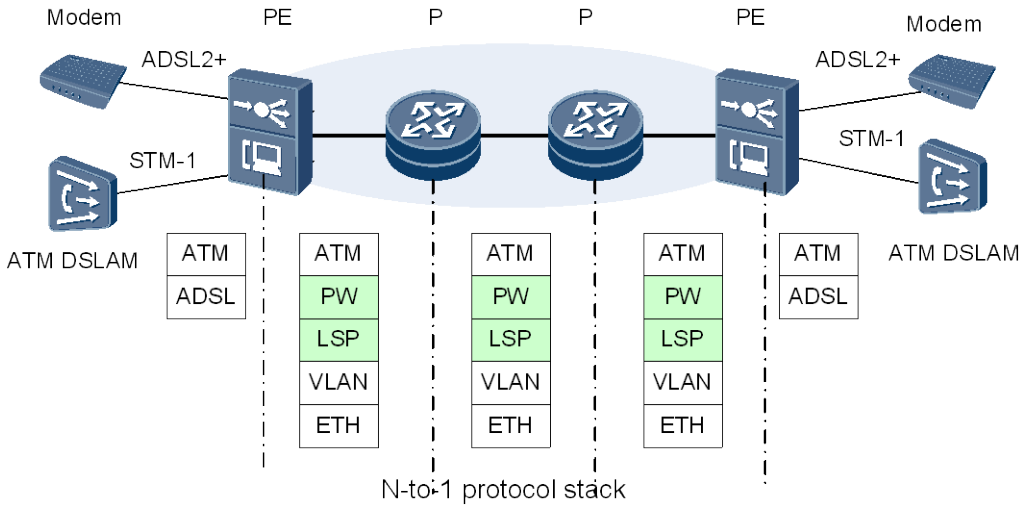


ATM PWE3 Principle

Processing of ATM PWE3 PDUs

Figure 12-17 shows how ATM pseudo wire emulation edge-to-edge (PWE3) protocol data units (PDUs) are processed on the provider edge (PE) and provider (P) devices. The MA5600T/MA5603T/MA5608T can function as a PE or a P device. The PE establishes an MPLS-based PW tunnel and encapsulates the original data packets (ATM cells or Ethernet packets) with two labels at the transmit (Tx) end for transmission. The P device forwards the packets. The PE at the receive (Rx) end decapsulates the received MPLS packets, restores the original data packets, and transmits the packets to users.

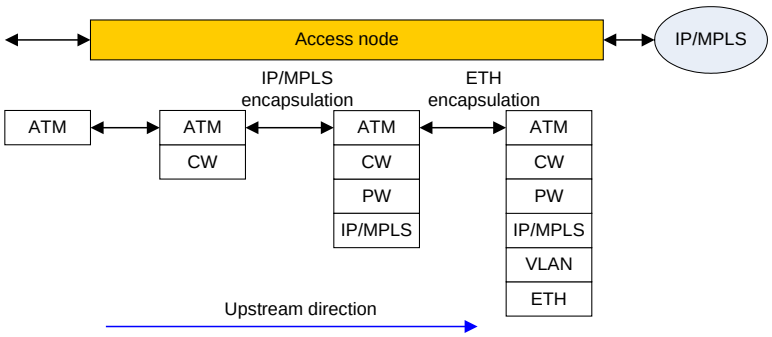
Figure 12-17 Processing of ATM PWE3 PDUs



Processing of ATM PWE3 Service

Figure 12-18 shows how the ATM PWE3 service (MPLS over IP or MPLS over MPLS encapsulation) is processed.

Figure 12-18 Processing of ATM PWE3 service



- Encapsulation/Decapsulation of ATM cells

  A control word is added to an ATM cell. Figure 12-19 shows the format of the control word.

  Figure 12-19 Format of the control word for an ATM cell

- MPLS over IP encapsulation (MPLS over MPLS encapsulation is the same.)

  The IP header and the PW label are added/deleted, and the PW label is mapped to a permanent virtual connection (PVC).

  - The outer IP header is used for transmitting the packet through the IP network, and the PW label is used for mapping to a PVC.
  - The source IP address in the IP header is the IP address of the Layer 3 interface, and the destination IP address is the IP address of the peer end (identical to the peer-address configured in the PW template).
  - The ToS bit needs to be configured by users. The value of the protocol field is 137 (identifying an MPLS unicast packet), the value of the DF bit is 1, and that of the MF bit is 0.
- Ethernet processing

  The upstream VLAN of the ATM PWE3 PDU is a service VLAN, which is the VLAN of the corresponding upstream port.

  The MAC address of the Layer 3 interface is filled in as the source MAC address of the ATM PWE3 upstream PDU, and the MAC address of the next-hop interface (this MAC address can be learned only by ARP) is used as the destination MAC address.

ATM PW N-to-1 (N > 1) Encapsulation

The following encapsulation formats are defined in the RFC4717 for the PWE3 emulation of ATM services in a packet-switched network (PSN). The MA5600T/MA5603T/MA5608T supports N-to-1 ($N \geq 1$) and ATM adaptation layer 5 (AAL5) service data unit (SDU) encapsulation.

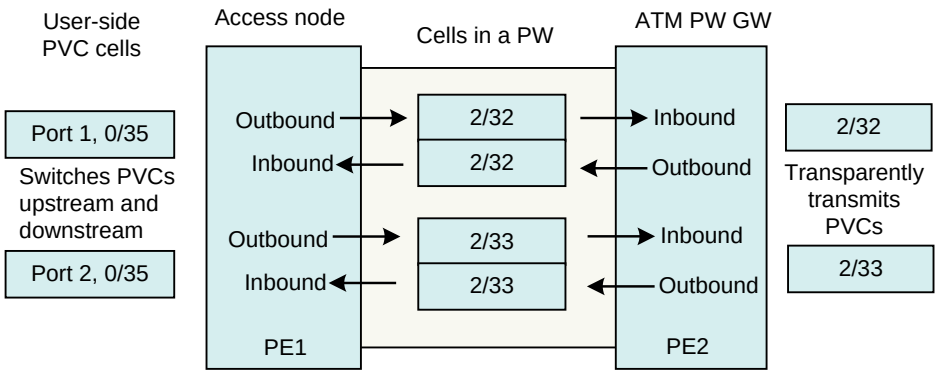
- N-to-1 (N > 1): Multiple ATM virtual channel connections (VCCs) or virtual path connections (VPCs) are transported in one PW.
- 1-to-1: Only one ATM VCC or VPC is transported in one PW.
- AAL5 SDU: Only the AAL5 CPCS-SDU payload is transported.
- AAL5 PDU: The AAL5 PDU, together with the PAD and CPCS-PDU, is transported.

For the N-to-1 (N > 1) ATM PWE3, user cells in multiple user-side PVCs are encapsulated into one PW. In order for the Rx end to differentiate these cells, the ATM service payload in a PW needs to contain the VPI/VCI information about the cells and the VPI/VCI information about each cell needs to be unique. During user service provisioning by carriers, however, the VPI/VCI values of all user PVCs are the same. As such, VPI/VCI switching is required when user PVCs are encapsulated into a PW to ensure unique VPI/VCI values for cell differentiation.

The MA5600T/MA5603T/MA5608T supports the following two PVC (VPI/VCI) switching methods.

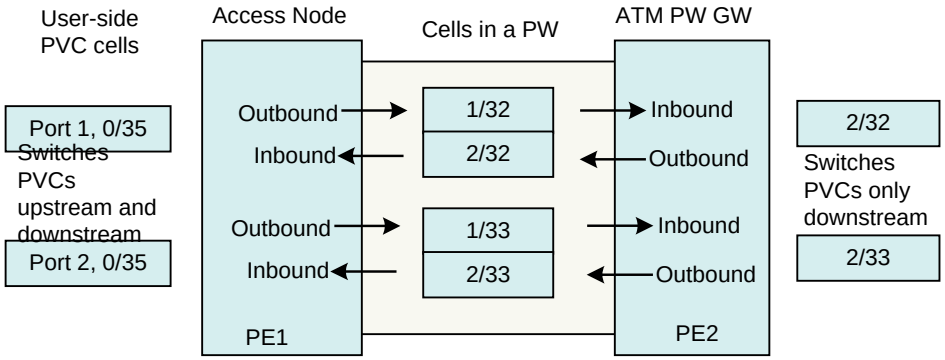
- The peer PE transparently transmits PVCs and does not perform downstream PVC switching, as shown in Figure 12-20.
  - In the upstream direction, PVC switching is performed on ingress PE1 for the cells with the same VPI/VCI values from different ports. This is to ensure the uniqueness of VPI/VCI values in a PW. After the cells are transmitted upstream to egress PE2, egress PE2 transparently transmits the cells without downstream PVC switching.
  - In the downstream direction, ingress PE2 does not perform PVC switching in the inbound direction but transparently transmits the cells to egress PE1. Egress PE1 then performs PVC switching and transmits the cells to the AC-side xDSL port.

Figure 12-20 PVC switching method 1: transparently transmitting PVC by the peer PE



- The peer PE performs downstream PVC switching, as shown in Figure 12-21.
  - In the upstream direction, PVC switching is performed on ingress PE1 with the same VPI/VCI values from different ports. This is to ensure the uniqueness of VPI/VCI values in a PW. After the cells are transmitted upstream to egress PE2, egress PE2 performs downstream PVC switching.
  - In the downstream direction, ingress PE2 does not perform PVC switching in the inbound direction but transparently transmits the cells to egress PE1. Egress PE1 then performs PVC switching and transmits the cells to the AC-side xDSL port.

Figure 12-21 PVC switching method 2: downstream PVC switching by the peer PE

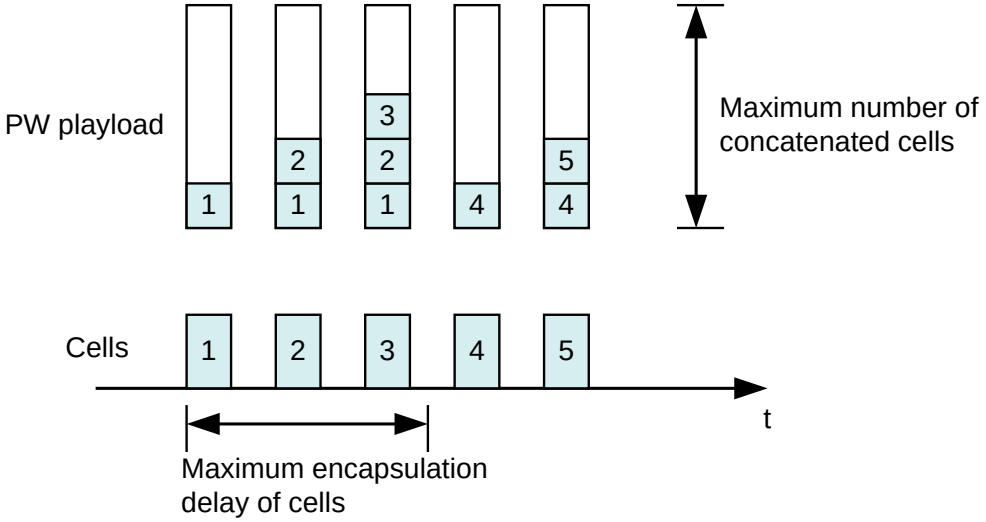


ATM Cell Concatenation

In order to improve transport efficiency on the PSN, multiple ATM cells are encapsulated in a single PW PDU. This process is called ATM cell concatenation.

With cell concatenation, cell transfer delay and jitter in the link are increased although the encapsulation efficiency is improved. For example, the more the cells are concatenated, the greater the delay of sending encapsulated cells. The MA5600T/MA5603T/MA5608T provides two parameters, maximum number of concatenated cells and maximum encapsulation delay of cells, for balance. As shown in Figure 12-22, during encapsulation of the concatenated cells, the cells are sent once the maximum encapsulation delay of cells or the maximum number of concatenated cells is reached.

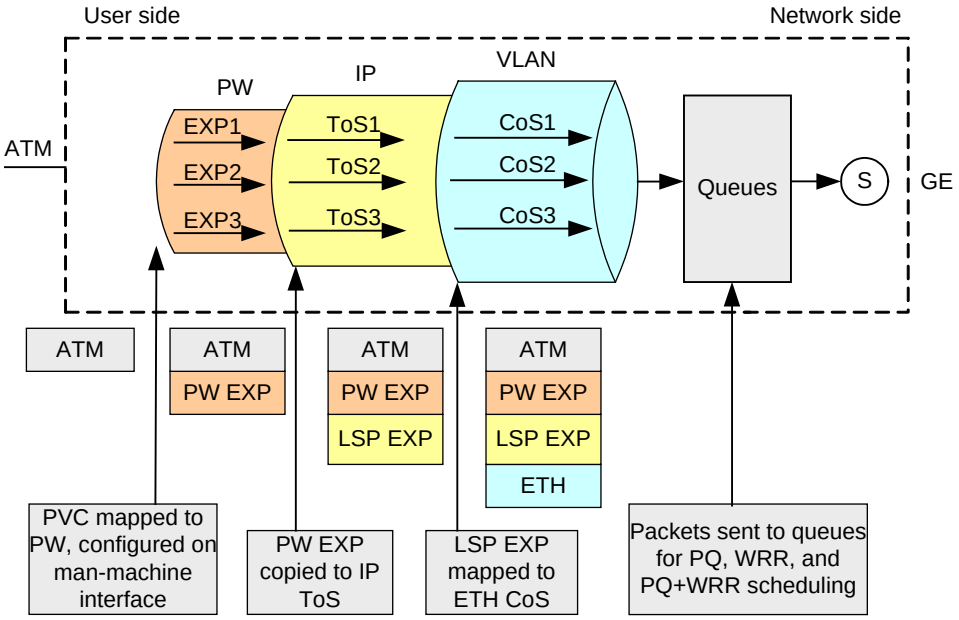
Figure 12-22 Maximum number of concatenated cells and maximum encapsulation delay of cells



QoS Processing of ATM PWE3 Service

Figure 12-23 shows the QoS processing of the ATM PWE3 service (MPLS over MPLS encapsulation).

Figure 12-23 QoS processing of upstream ATM PWE3 service (MPLS over MPLS encapsulation)



PW-based Dual-Bucket CAR

QoS is required for the user ATM cells carried in ATM PWE3 over the PSN network. Due to mechanism differences, the ATM traffic policing mechanism needs to map the MPLS traffic policing mechanism on an ingress PE, and a reverse mapping is required on an egress PE.

In the upstream direction of an ingress PE, PW-based dual-bucket CAR, that is, two rate three color marker (trTCM), is performed according to PW CAR or LSP CAR. With this mechanism, ATM cells whose rate is lower than committed information rate (CIR) are marked with the default CoS value of ATM over Ethernet (AoE) traffic streams, and ATM cells whose rate is higher than CIR and lower than peak information rate (PIR) are re-marked with a low-priority CoS value, while ATM cells whose rate is higher than PIR are dropped. During encapsulation of PW PDUs, the CoS of the AoE packet is mapped to the EXP field of the inner PW label and then to the EXP field of the outer MPLS label. Then, traffic policing is performed over the PSN network based on the EXP field of the outer MPLS label, as shown in Figure 12-24.

In the downstream direction of an egress PE, the egress PE works with the ingress PE to put the ATM cells carrying the default CoS tag and those carrying the low-priority CoS tag into the same queue according to the CoS information carried in the EXP field. Also, different early drop thresholds are configured for these two CoS tags to ensure that ATM cells whose rate is lower than CIR have a higher priority when congestion occurs.

In the upstream direction, the MA5600T/MA5603T/MA5608T implements PW-based dual-bucket CAR on the SPUB board, which achieves trTCM by CoS re-marking based on CIR and PIR, as shown in Figure 12-24. In the downstream direction, the MA5600T/MA5603T/MA5608T does not perform CAR or CoS-based tail drop on the SPUB board, but implements queue-based early drop on the xDSL board according to the CoS early drop threshold, as shown in Figure 12-25.

Figure 12-24 PW-based dual-bucket CAR on the SPUB board (in the upstream direction)

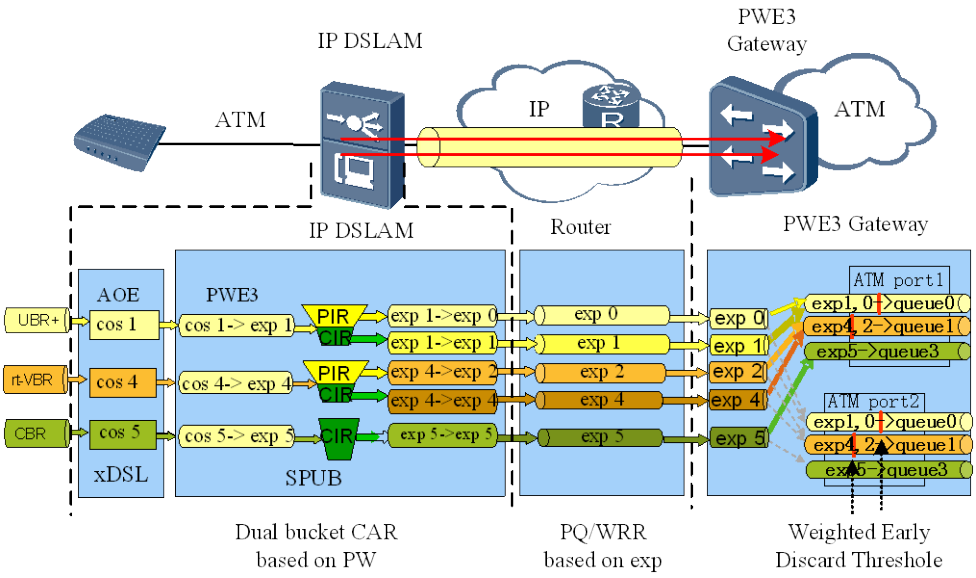
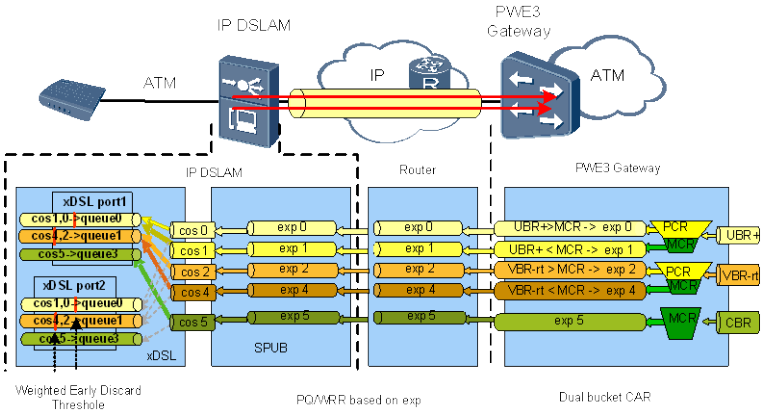


Figure 12-25 Threshold-based early drop on the xDSL board (in the downstream direction)



When cell concatenation is used for binding of ATM cells to PWs, PVCs of the same type are generally bound to the same PW. Because the AoE CoS value of the ATM cells is copied as the CoS value of the PW, the AoE CoS priority is affected when cells of different priorities are concatenated.

- When PW does not use cell concatenation, a PW PDU contains only one ATM cell. In this case, the AoE CoS value of the ATM cell is directly copied as the EXP value of the PW PDU.
- When PW uses cell concatenation, a PW PDU contains multiple ATM cells. In this case, if the AoE CoS values of these ATM cells are different, the CoS value indicating the highest priority will be copied as the EXP value of the PW PDU. Then, the AoE CoS values (equaling the EXP value of the PW PDU) of ATM cells in the same PW PDU will be the same in the downstream direction, which affects queue scheduling on the xDSL board.

Principle of ETH PWE3

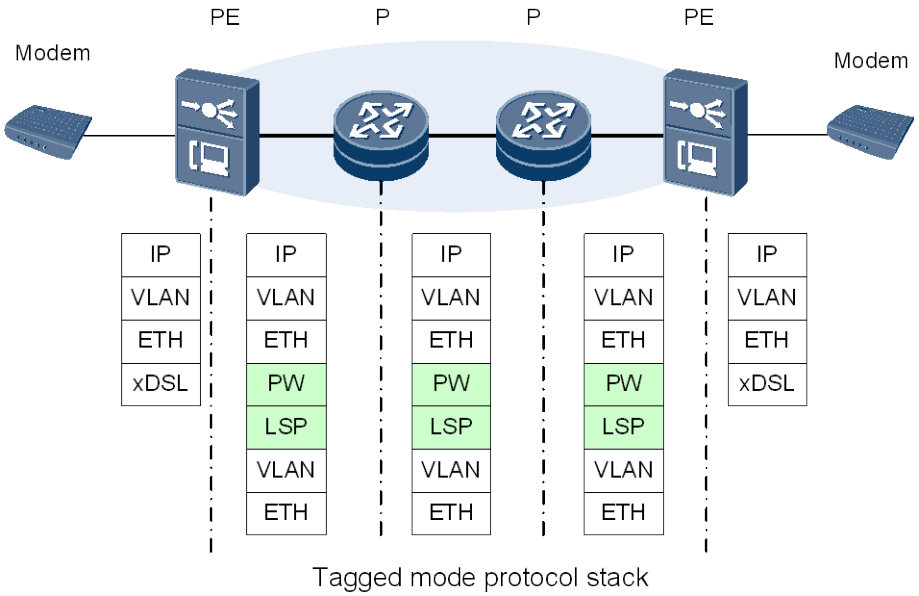
Processing Flow of ETH PWE3 Protocol Packets

Two encapsulation formats are defined in RFC4448 for the PWE3 emulation of Ethernet service in a PSN network.

- Tagged mode. In this mode, the packet going upstream carries the PW VLAN tag in the payload, and is stripped of the PW VLAN tag when going downstream.
- Raw mode. In this mode, the packet going upstream does not carry the PW VLAN tag; the PW payload, however, can carry the service VLAN tag.

Figure 12-26 shows the processing flow of ETH PWE3 protocol packets on the PE and P devices. The MA5600T/MA5603T/MA5608T can function as a PE or a P device. The PE establishes an MPLS-based PW tunnel, encapsulates the user data packets with two labels at the Tx end and transmits the packets. The P device forwards the packets. The PE at the Rx end decapsulates the received MPLS packets, restores the original user data packets, and transmits the packets to the user.

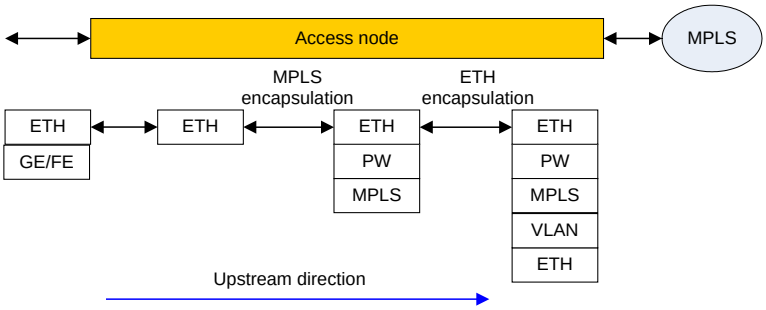
Figure 12-26 Processing flow of ETH PWE3 protocol packets



Processing Flow of ETH PWE3 Service

Figure 12-27 shows the processing flow of ETH PWE3 service (MPLS over MPLS encapsulation).

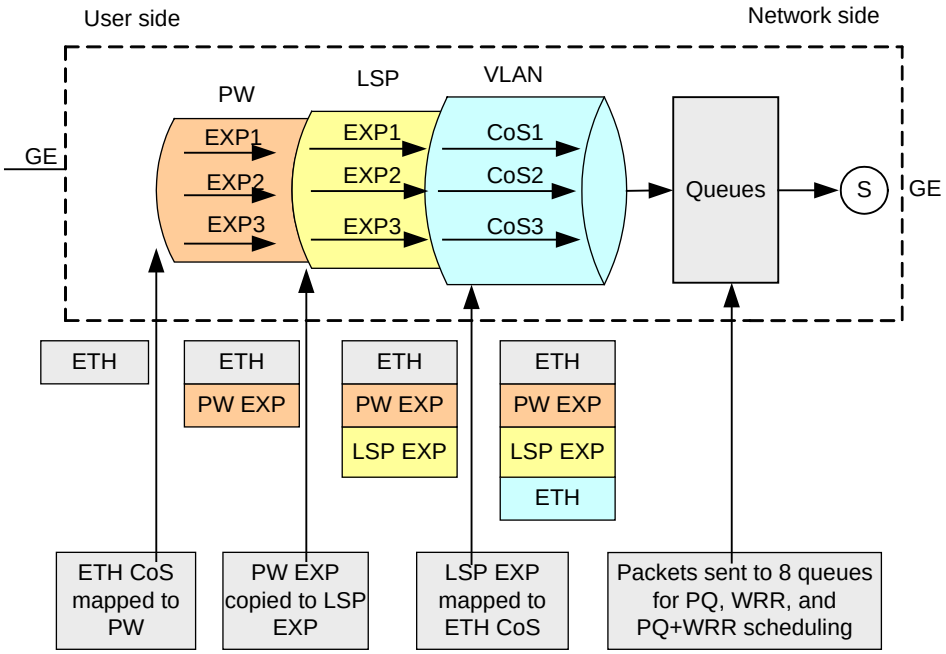
Figure 12-27 Processing flow of ETH PWE3 service



QoS Processing Flow of ETH PWE3 Service

Figure 12-28 shows the QoS processing flow of ETH PWE3 service (MPLS over MPLS encapsulation).

Figure 12-28 QoS processing flow of upstream ETH PWE3 service (MPLS over MPLS encapsulation)



Traffic Label Principle

Context

As services are developing, service traffic over a PW becomes heavier and heavier. For example, in the case of the wholesale service, certain carriers encapsulate all service traffic on hundreds or thousands of xDSL ports into a PW, and as a result traffic over the PW reaches the gigabit level. As such, a PW carrying such a heavy traffic is called fat PW. The fat PW burdens the equipment that it traverses, and results in congestion, packet loss, and unguaranteed QoS.

To solve problems incurred by the fat PW, IETF proposed a traffic label solution: Traffic over a fat PW takes different paths from the PW ingress PE to the PW egress PE within the network through load balancing (ECMP). To achieve PWE3 load balancing, PW data at the PW ingress PE are segmented into bundles of data streams and an MPLS label (traffic label) is allocated to every data stream. In this way, every traffic label identifies a different data stream and the traffic label is stored in the innermost area of the ingress PE label stack. Later, data is forwarded in the load balancing mode according to the traffic label, taking different data streams along different paths within the network.

In terms of PW load balancing, data over a PW arrives at the destination along different paths and this may incur disordered packets. Given this, this technology is applicable to only those services that are not sensitive to disordered packets, such as the Ethernet service. For the ATM and TDM emulation services, this technology is not applied because they have strict requirements on timing and order of packets.

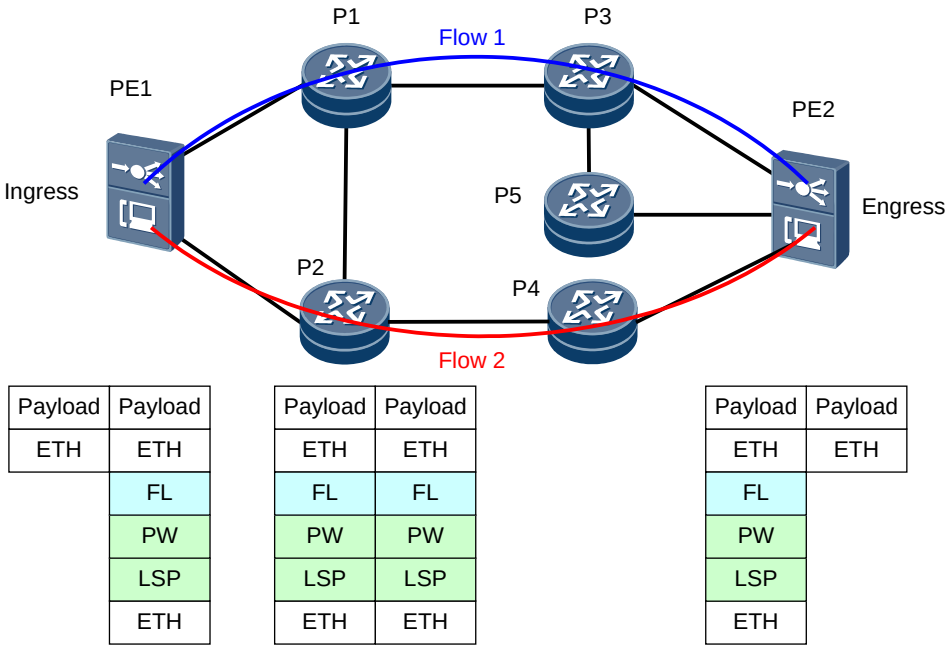
PW load balancing is implemented based on equal cost LSP that is based on the equal cost route.

Application Description

Figure 12-29 shows an application over an existing network where a large amount of the legacy equipment does not support the traffic label. In this application, the PE supports generating the traffic label and performing load balancing but the P has no traffic label capability (for example, the legacy equipment over the network).

- PE1 generates traffic label (FL) and at the same time performs load balancing (flow1 and flow2). PE2 removes the FL.
- P1, P2, and P3 do not support traffic label for load balancing and they only forward data like a common P.

Figure 12-29 Traffic label application (P equipment does not support traffic label for load balancing)



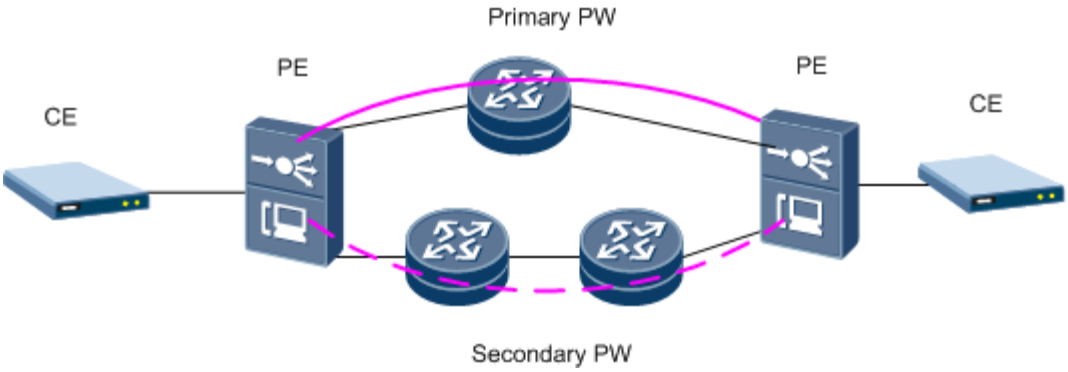
PW 1:1 Redundancy

PW redundancy is used for the PW protection switching, which has a similar function as the MPLS OAM. Different from the MPLS OAM (protecting the outer tunnel), PW redundancy protects the inner tunnel.

Through PW redundancy, data is switched to the standby PW if the active PW is faulty (such as the LDP session is down, the tunnel is deleted, the protocol communication is faulty, the route status is changed, or there is no VCCV response). In this case, the original standby PW becomes the active PW.

Figure 12-30 shows the MA5600T/MA5603T/MA5608T supports the following PW 1:1 redundancy applications.

Figure 12-30 PW 1:1 redundancy applications



PW Redundancy

PW Redundancy Signaling Mechanism

Introduction of the PW protection mechanism will break the original model of 1-to-1 mapping between AC and PW in PWE3. To keep the original forwarding action, you must ensure that only one PW in the redundancy PW group is in the active state and other PWs are in the standby state.

The LDP PW signaling (RFC4447) requires to use PW status TLV to transfer the PW forwarding status. PW status TLV can be carried by the label mapping message or notification message. PW status TLV is a 32-bit status code and each bit identifies a PW forwarding state. Based on this status code, PW redundancy introduces a new PW status code (0x00000020 - PW forwarding standby) to indicate that the PW is in the standby state currently.

Primary/Secondary and Active/Inactive

There are two couples of important concepts in PW redundancy and the detailed descriptions are as follows:

- Primary/Secondary refers to the PW forwarding priority and is the PW configuration parameter.

  The primary PW is preferentially used to forward traffic and the secondary PW is used to protect primary PW. The primary PW is used to forward traffic when the state of the primary and secondary PWs is the same. Currently, only one secondary PW can be configured for each primary PW.
- Active/Inactive refers to the PW forwarding status. It indicates the PW running status and is not the configuration parameter.

  Only the PW in the active state can be used to forward traffic. The local active or inactive state of a PW is determined by the local and remote signaling status and priority (configured primary/secondary) of the PW. Only the PW in the optimal state and with the highest priority can be selected as the active PW to forward traffic, and all other PWs are in the inactive state. PWs in the inactive state are not used to forward traffic but can be enabled to receive traffic (can be used only for VLL PW).

PW Redundancy Working Mode

The PW redundancy working mode is specified on PE that is configured with active and standby PWs.

Master/Slave mode:

In this mode, the local end determines the active or standby state of the PW and uses the signaling protocol to notify the remote end; the remote PE can sense the active or standby state. The active/standby relationship on the PW side and the active/standby relationship on the AC side do not affect each other. Therefore, faults can be isolated between the PW side and the AC side. The MA5600T/MA5603T/MA5608T currently can function as a master in this mode.

Independent mode:

In this mode, the active or standby state of the local PW is determined by the negotiation result of the remote AC side; the remote end notifies the local end of the active or standby state. The protection switching due to faults on the AC side will cause the protection switching on the PW side. Therefore, faults cannot be isolated.

PW OAM (VCCV)

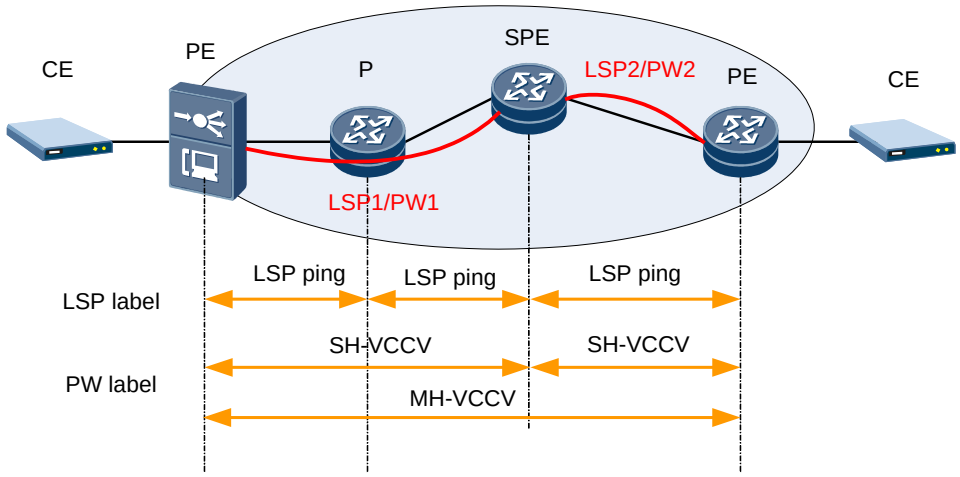
VCCV Ping Application

Virtual Circuit connectivity verification (VCCV) is an end-to-end mechanism to detect and diagnose PW faults. To put it simply, VCCV is a control channel for transmitting connectivity verification messages between PW ingress and PW egress.

VCCV ping is a tool for manually detecting the connectivity status of virtual circuits. It is implemented through extending LSP ping. VCCV defines a series of messages exchanged among PEs to verify PW connectivity. To ensure that the VCCV packet traverses the same path as the data packet in the PW, the VCCV packet must be encapsulated in the same encapsulation mode as the PW and must traverse the same tunnel as the PW packet.

VCCV ping can be used on the U-PE to detect PW connectivity, including detecting the connectivity of the static PW, dynamic PW, single-hop PW, and multi-hop PW. The MA5600T/MA5603T/MA5608T supports single-hop-VCCV (SH-VCCV) ping and does not support multi-hop-VCCV (MH-VCCV) ping temporarily. Figure 12-31 shows the VCCV ping application.

Figure 12-31 VCCV ping application



Principle

VCCV ping is implemented through the VCCV packet and the LSP ping packet therein carries the target FEC stack. Figure 12-32 describes the parameters of the VCCV packet.

Figure 12-32 VCCV packet parameters

0x0c	0x04	CC types	CV types

- CC Types indicates the control channel type. Figure 12-33 describes the CC for the VCCV function, which is defined in RFC5085.

  Figure 12-33 CC in the VCCV packet

```
MPLS Control Channel (CC) Types:

Bit (Value)     Description
============    ============================================
Bit 0 (0x01) - Type 1: PWE3 Control Word with 0001b as
               first nibble (PW-ACH, see [RFC4385])
Bit 1 (0x02) - Type 2: MPLS Router Alert Label
Bit 2 (0x04) - Type 3: MPLS PW Label with TTL == 1
Bit 3 (0x08) - Reserved
Bit 4 (0x10) - Reserved
Bit 5 (0x20) - Reserved
Bit 6 (0x40) - Reserved
Bit 7 (0x80) - Reserved
```

  - Type1: control word channel. Whether VCCV is performed depends on the control word (0001 or not). SH-VCCV and MH-VCCV are supported. If the PE supports control word, CC type1 is used preferentially.
  - Type2: MPLS router alert channel. Whether VCCV is performed depends on a specific label value (label = 2). SH-VCCV is supported and MH-VCCV is not supported.

- Type3: maximum-hop channel. VCCV is performed if the TTL value of the inner label of MPLS is 1. SH-VCCV and MH-VCCV are supported.

- CV Types indicates the connectivity verification type. Figure 12-34 describes the CV for the VCCV function, which is defined in RFC5085. The MA5600T/MA5603T/MA5608T supports only CV of the LSP ping type.

Figure 12-34 CV in the VCCV packet

```
MPLS Connectivity Verification (CV) Types:

Bit (Value)     Description
============    ========================================
Bit 0 (0x01) - ICMP Ping
Bit 1 (0x02) - LSP Ping
Bit 2 (0x04) - Reserved
Bit 3 (0x08) - Reserved
Bit 4 (0x10) - Reserved
Bit 5 (0x20) - Reserved
```

T-PE peers at both sides negotiate CC and CV capabilities during PW set-up, and then send the same CC and CV types used by VCCV ping as the negotiation result. If the PE supports control word, CC type1 is used preferentially. SH-VCCV and MH-VCCV are implemented through setting different inner PW label TTL values.

Figure 12-35 shows the CC Type1 VCCV flow.

Figure 12-35 CC Type1 MH-VCCV/SH-VCCV flow

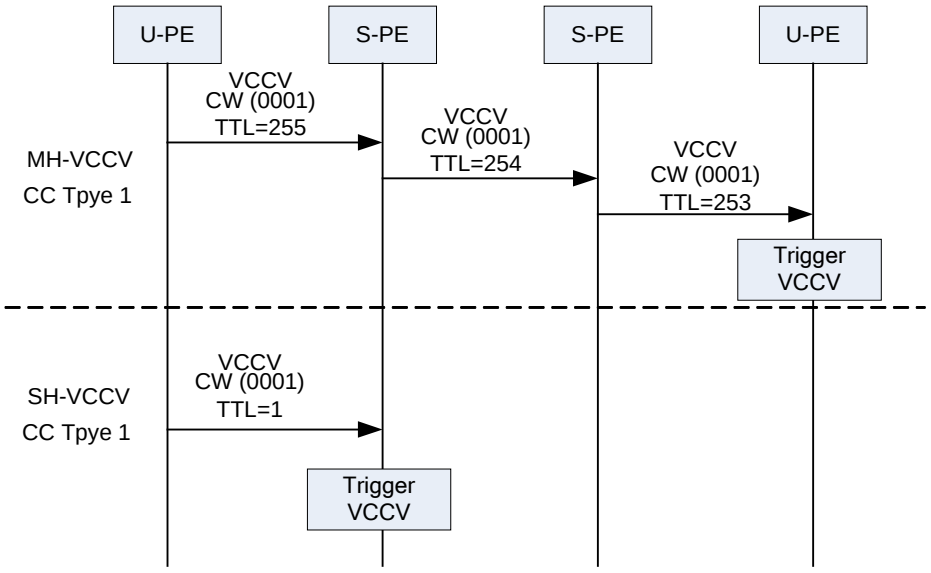
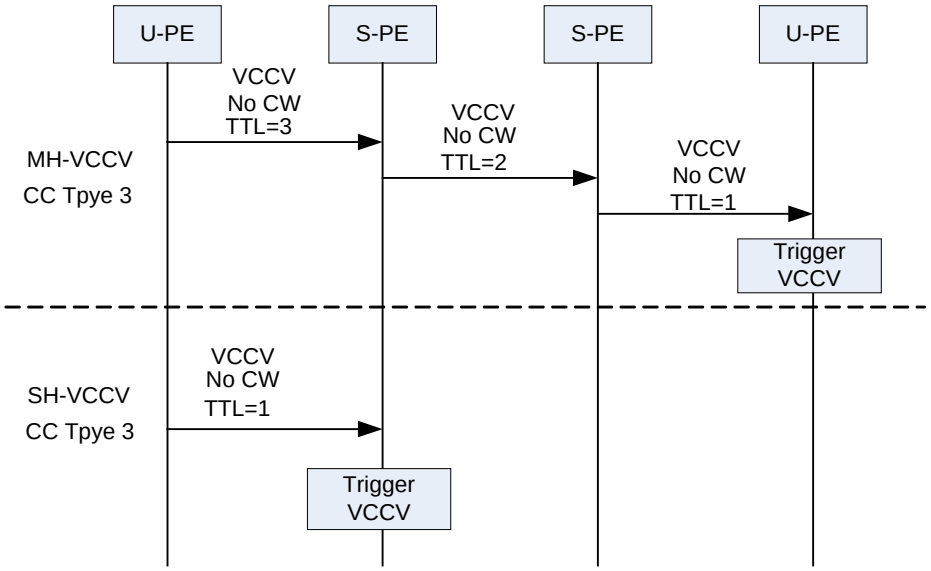


Figure 12-36 shows the CC Type3 VCCV flow.

Figure 12-36 CC Type3 MH-VCCV/SH-VCCV flow



The U-PE and S-PE differ in processing the VCCV packet:

- The S-PE pays attention to only the PW TTL value. If the PW TTL value is 1, VCCV is performed.
- The U-PE pays attention to not only CW (VCCV is performed if the first half-byte is 0001) but the PW TTL value (VCCV is performed if the PW TTL value is 1).
- If the U-PE initiates VCCV to the S-PE, a proper TTL value needs to be set, irrespective of whether CW is used.

12.1.4 Network Applications

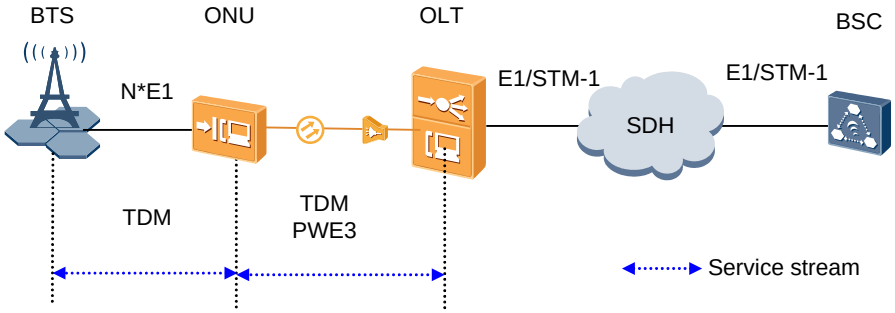
Network Applications of TDM PWE3

For details on the corresponding clock solution of the TDM service, see 32.5 Physical Layer Clock/Time Synchronization.

- Network application: converting the TDM PWE3 service or the native TDM service into the E1/STM-1 service for upstream transmission

  As shown in Figure 12-37, the mobile 2G base station is connected to the ONU through TDM E1. The ONU performs TDM PWE3 emulation, or the ONT encapsulates the TDM frame into the GPON GEM frame directly to transmit TDM service over the GPON network (native TDM). The OLT terminates the TDM PWE3 signals and transmits the signals to the upstream SDH network through the E1 or STM-1 port.

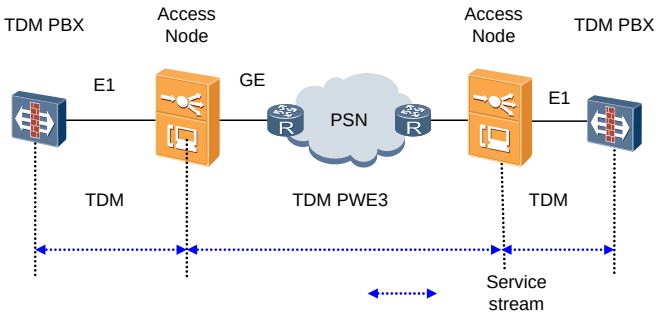
Figure 12-37 Network application: converting the TDM PWE3 service into the E1/STM-1 service for upstream transmission



- Enterprise Private Line Service Using TDM PWE3

  As shown in Figure 12-38, an enterprise TDM PBX is connected to the MA5600T/MA5603T/MA5608T through the E1 port. The MA5600T/MA5603T/MA5608T performs the TDM PWE3 emulation and transmits service streams to the peer TDM PWE3 device through the packet switched network (PSN). In this way, the enterprise private line service is implemented through TDM PWE3.

Figure 12-38 Networking application of enterprise private line service using TDM PWE3



- E1/T1 Integrated Access

  Figure 12-39 shows the E1/T1 integrated access.

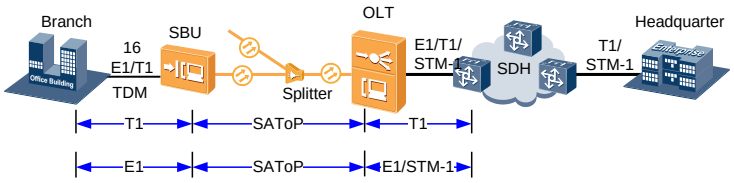
  An ONU is connected to an enterprise PBX through the TDM E1/T1 port. The ONU services are sent upstream to the OLT in SAToP mode and then to the SDH network.

  The MA5612 or MA5628 can be the ONU.

  E1/T1 integrated access achieves the following values to customers:

  - Legacy SDH network resources are fully unitized.
  - The interfaces are unified, simplifying deployment and maintenance.

Figure 12-39 E1/T1 integrated access



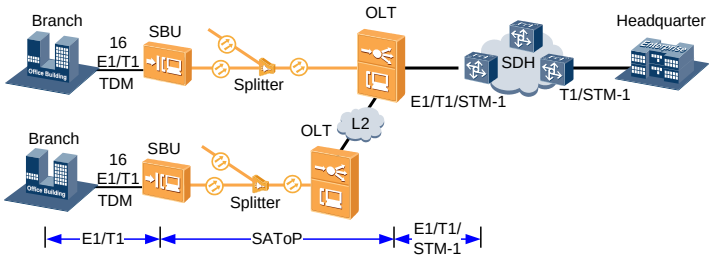
- SAToP Service Termination on the OLT of Another Node

  Figure 12-40 shows the termination of SAToP services on the OLTs of another node. Carrier SDH networks are gradually migrated. In a city, services of only some nodes can be sent upstream through STM-1 ports, and the other nodes need to be connected to regions that have SDH resources using GE/10GE ports.

  In this case, an SBU is connected to an enterprise PBX through the TDM E1/T1 port. The SBU services are sent upstream to an OLT in SAToP mode. The OLT then transparently transmits the SBU services to another OLT through the GE/10GE port. It is recommended that the two OLTs are directly connected. The two OLTs can also be interconnected with each other through the upper-layer convergence switch after passing at most two hops (it is recommended that services do not traverse the third-party IP network). Then, the services are sent upstream to the SDH network. In this way, enterprises that do not have SDH resources due to region restrictions can successfully deploy enterprise private line access services.

  The MA5612 or MA5628 can be the SBU.

Figure 12-40 SAToP service termination on the OLT of another node

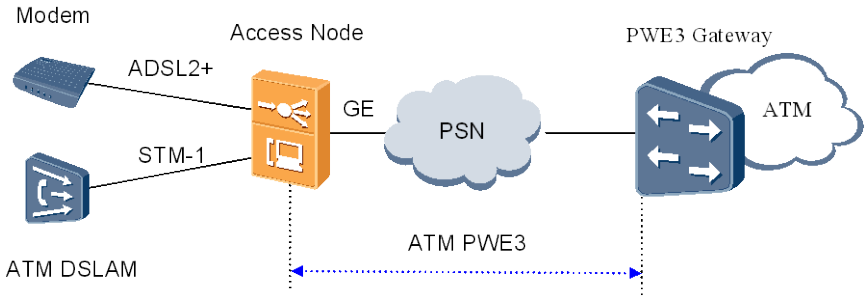


Network Applications of ATM PWE3

- Network application: implementing PWE3 private line upstream transmission in ATM access

  As shown in Figure 12-41, when the MA5600T/MA5603T/MA5608T is connected to the ATM DSLAM or ADSL2+ modem, ATM private line service can be implemented between the MA5600T/MA5603T/MA5608T and the peer ATM BRAS through ATM PWE3 private line. The ATM PWE3 private line service is applicable to ATM network restructuring.

Figure 12-41 Network application: implementing PWE3 private line upstream transmission in ATM access

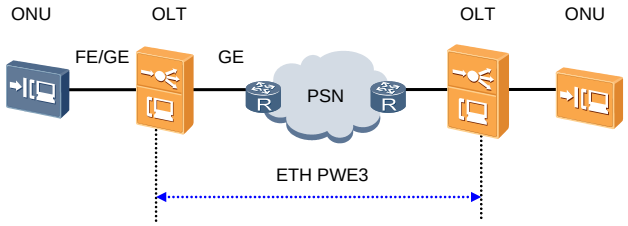


Network Applications of ETH PWE3

- Network application: implementing PWE3 private line upstream transmission in FE/GE access

  As shown in Figure 12-42, the enterprise router is connected to the OLT through FE/GE. The OLT interconnects with the peer ETH PWE3 device of the enterprise through the ETH PWE3 private line to implement the ETH private line service.

Figure 12-42 Network application: implementing PWE3 private line upstream transmission in FE/GE access



12.2 Native TDM

In Native TDM, TDM frames are directly encapsulated to GPON GEM frames in TDMoGEM mode. This mode features simple encapsulation, small network cost, and guaranteed link quality.

12.2.1 Introduction

Definition

By using the standard 8 kHz (125μm) frames, the GPON GTC layer is synchronous in nature. Therefore, GPON can support the TDM service. This is called Native TDM.

In Native TDM, TDM frames are directly encapsulated to GPON GEM frames in TDMoGEM mode. This mode features simple encapsulation, small network cost, and guaranteed link quality.

Purpose

Currently, the circuit switched network is evolving to the packet switched network. During the deployment of the packet switched network, the method to provide traditional circuit switching service over the packet switched network must be taken into consideration. In a GPON network deployment, the traditional TDM service can be delivered over the PSTN network through the Native TDM mode.

12.2.2 Reference

The following lists the reference documents of Native TDM:

- ITU-T G.984.1 General characteristics for Gigabit-capable Passive Optical Networks (GPON)
- ITU-T G.984.2 Gigabit-capable Passive Optical Networks (GPON): Physical Media Dependent (PMD) layer specification
- ITU-T G.984.3 Gigabit-capable Passive Optical Networks (GPON): Transmission convergence layer
- ITU-T G.984.4 Gigabit-capable Passive Optical Networks (GPON): ONT management and control interface specification

12.2.3 Principle

Basic Principle

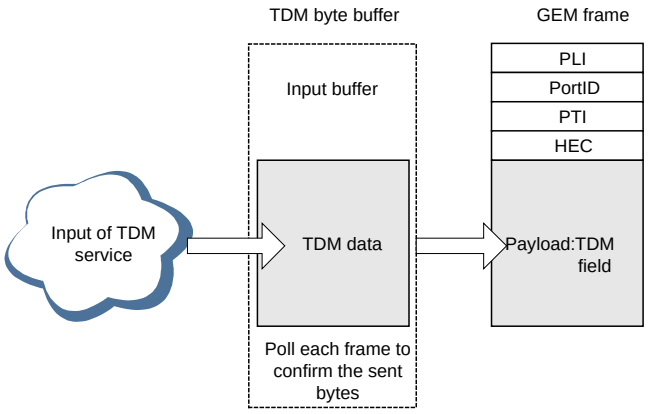
The MA5600T/MA5603T/MA5608T supports E1 and STM-1 upstream transmission mode. The following section uses E1 upstream transmission as an example.

NOTE

The application of STM-1 upstream transmission is similar to the application of E1 upstream transmission. The only difference lies in that different daughter boards are attached to the TOPA boards in two applications. When the TOPA board provides STM-1 ports, STM-1 frames are generated. One STM-1 frame can encapsulate and multiplex up to 63-channel E1 service signals (One STM-1 frame contains one VC4, one VC4 maps 63 VC12s, and one VC12 maps one-channel E1 service signal).

In the TDMoGEM mode, the TDM frame is directly encapsulated to the GPON GEM frame so that the TDM service can be transmitted in a GPON network. Figure 12-43 shows the basic principle of Native TDM.

Figure 12-43 Basic principle of Native TDM



Mapping of the TDM service to the GEM frames allows variation of the GEM frame length based on frequency offset of the TDM service. The length of the TDM field is specified by the PLI field.

In the adaptation process of the TDM source, the input data enters a queue in the input buffer. When a frame arrives (namely for every 125 μs) there, the multiplexing entity of the GEM frame records the number of bytes to be sent in the current GEM frame. In general, the PLI specifies a fixed byte count based on the TDM nominal rate. But it happens that the bytes larger or smaller than the fixed byte count are sent, which are reflected in the PLI field.

The TDMoGEM mode does not distinguish whether it is a structured service such as the voice and PBX access services, or an un-structured service such as private line service. All services are handled as un-structured services. GPON access transparently delivers only the E1 service without performing the refined service processing. In other words, GPON access provides the long-haul E1 transmission.

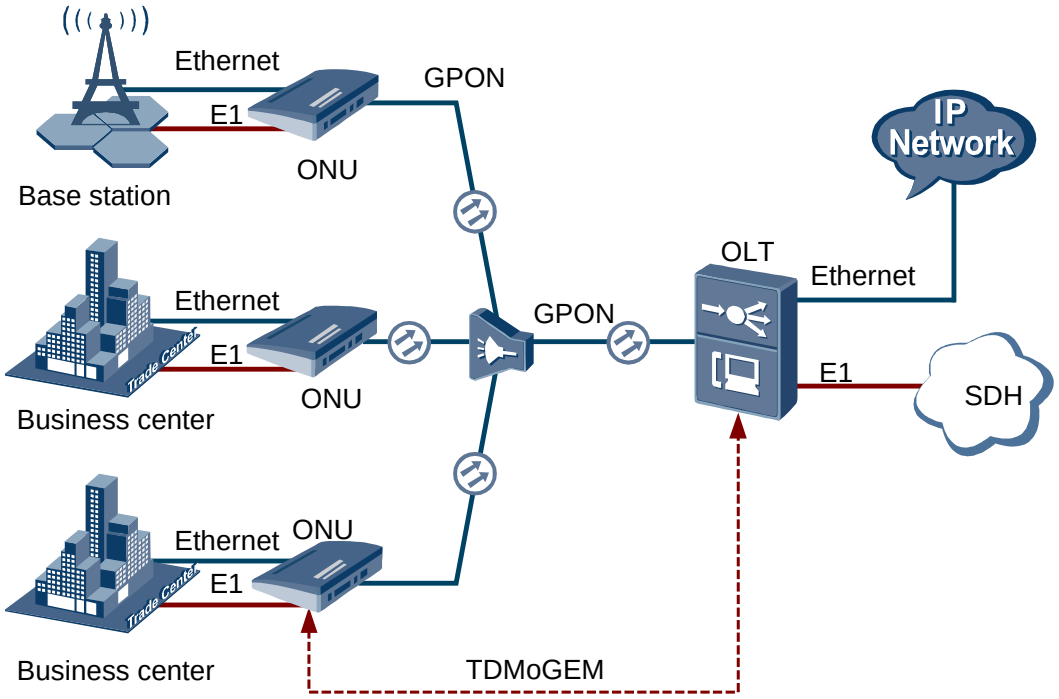
In general, the private line service has an independent clock, which is asynchronous with the GPON line clock. The jitter and frequency difference exist between the two clocks. Therefore, the E1 service traffic can be transmitted only after the E1 rate adjustment occurs at the ONU.

The bit asynchronous mapping mode of SDH is adopted while the GPON line clock or system clock is used as the reference clock. Through the bit adjustment, the E1 service traffic is mapped to the payload section of the tributary unit to form a standard VC12 virtual channel, which is delivered in the GEM frame. At the receiver end, de-byte adjustment is adopted to recover the original clock and the E1 service traffic.

Application

Figure 12-44 shows the TDMoGEM network application. The E1 ports of the ONU/ONT access the TDM traffic from the base stations and enterprises. The ONU/ONT sends both the Ethernet traffic and the TDM traffic to the OLT. The OLT differentiates the Ethernet traffic and the TDM traffic, and sends the traffic to the IP network and the SDH network, respectively.

Figure 12-44 Native TDM network application



12.3 Configuring the PWE3 Private Line Service

Pseudo wire emulation edge-to-edge (PWE3) uses LDP or RSVP-TE as the signaling protocol and carries various Layer 2 services of the customer edge (CE) over the MPLS LSP or TE tunnel, transparently transmitting the Layer 2 data of the CE.

PWE3 Service Model

According to the PWE3 service model, PWE3 is indicated by the outer packet switch network (PSN) tunnel label and the inner label (PW demultiplexer).

The PSN layer can select the MPLS or IP technology and the PW demultiplexer can select the MPLS, UDP, or layer-2 tunneling protocol (L2TP) technology. The PWE3 outer label and inner label support the following combinations: MPLS over MPLS, MPLS over IP, UDP over IP, and L2TP over IP. The MA5600T/MA5603T/MA5608T supports the first three.

Network Application

Figure 12-45 shows the network application of the MPLS PWE3.

As shown in the figure, the mainstream applications of the MPLS PWE3 supported by the MA5600T/MA5603T/MA5608T are as follows:

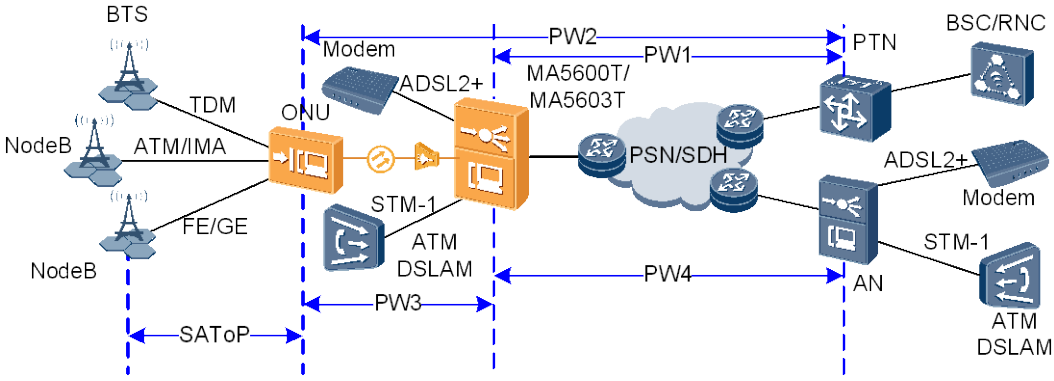
- TDM PWE3: A mobile 2G base station is connected to the ONU through the TDM E1 port. The ONU implements the TDM PWE3, transmitting traffic streams to the peer TDM PWE3 device through the PSN. The MA5600T/MA5603T/MA5608T functions as a Layer 2 transparent transmission device, PE device, or P device.

- ATM PWE3: The IMA service data of a 3G base station is connected to the ONU through the E1 port. The ONU restores the IMA service to the ATM service and encapsulates the ATM service on the ATM PWE3 private line for connecting to the peer ATM PWE3 device (PTN device in the figure). The MA5600T/MA5603T/MA5608T functions as a Layer 2 transparent transmission device or P device.
- ETH PWE3: A 3G base station is connected to the ONU through the FE/GE port. The ONU performs the ETH PWE3 encapsulation for interconnecting with the peer ETH PWE3 device. The MA5600T/MA5603T/MA5608T functions as a Layer 2 transparent transmission device or P device.

NOTE

- The MA5600T/MA5603T/MA5608T can function as a Layer 2 transparent transmission device or PE/P device, determined by the service requirement.
- As shown in the following figure, in PW1, PW3, and PW4, the MA5600T/MA5603T/MA5608T functions as a PE device that initiates or terminates the PW; in PW2, the MA5600T/MA5603T/MA5608T functions as a Layer 2 transparent transmission device or P device.

Figure 12-45 MPLS PWE3 network application when the MA5600T/MA5603T/MA5608T functions as a Layer 2 transparent transmission device, PE device, or P device



Procedure

According to the PWE3 service model, PWE3 configurations include the outer tunnel configuration, inner PW configuration, and tunnel protection. Therefore, the configuration procedure is as follows.

12.3.1 Configuring the PWE3 Outer Tunnel

To provide services across the IP network or MPLS network, the MA5600T/MA5603T/MA5608T supports PW over the IP tunnel or MPLS tunnel to transparently transmit services in the IP network.

Prerequisites

1. The loopback interface IP address must be configured.
2. The LSR ID must be configured.
3. The global MPLS and MPLS TE functions must be enabled.
4. The OSPF protocol must be successfully configured on each device in the network (the host route of each port must be successfully advertised).

Context

According to the upper-layer PSN type, namely MPLS network or IP network, the PWE3 outer tunnel is categorized as MPLS tunnel and IP tunnel.

Different PWE3s support different tunnel encapsulation formats. Pay attention to the following points during the configuration:

- TDM PWE3 supports the following PWE3 tunnel encapsulation formats: MPLS over MPLS and MPLS over IP
- ATM PWE3 supports the following PWE3 tunnel encapsulation formats: MPLS over MPLS and MPLS over IP.
- ETH PWE3 supports only the MPLS over MPLS encapsulation format.

Procedure

- Configure the MPLS TE tunnel.
  a. In the global config mode, run the **interface tunnel** command to create a tunnel interface and enter the tunnel interface mode.
  b. Run the **tunnel-protocol mpls te** command to configure the tunnel protocol to MPLS TE, that is, configure the tunnel interface to work in the TE tunnel mode.
  c. Run the **destination** *ip-address* command to configure the destination IP address of the tunnel. Generally, the LSR ID of the ingress is used.
  d. Run the **mpls te tunnel-id** command to configure the tunnel ID.
  e. Run the **mpls te signal-protocol** { **rsvp-te** | **static** } command to configure the signaling protocol for the MPLS TE tunnel.

     According to whether the MPLS TE tunnel uses the dynamic signaling protocol, the tunnel is categorized as static MPLS TE tunnel and MPLS RSVP-TE tunnel.

     - Static MPLS TE tunnel: The forwarding information and resource information are configured manually, and the signaling protocol and path calculation are not involved. Because the MPLS-related control packets are not exchanged, fewer resources are used. The static tunnel, however, cannot be dynamically adjusted according to network changes. Therefore, the actual application is limited.
     - MPLS RSVP-TE tunnel: MPLS TE creates the LSP tunnel along a specified path through RSVP-TE and reserves resources. Thus, carriers can accurately control the path that traffic traverses to avoid the node where congestion occurs. This solves the problem that certain paths are overloaded and other paths are idle, utilizing the current bandwidth resources sufficiently.

  f. (Optional) Run the **mpls te bandwidth** command to configure the bandwidth of the tunnel. After the configuration is completed, only the VLAN interface meeting this bandwidth requirement is selected as the node traversed by an MPLS TE tunnel when the MPLS TE tunnel is created.

     If the MPLS TE tunnel is only used to change the data transmission path, you may not configure the bandwidth of the tunnel.
  g. (Optional) Run the **mpls te path explicit-path** command to configure the explicit path used by the MPLS TE tunnel.

     To limit only the bandwidth of the MPLS TE tunnel but not the transmission path, you may not configure the explicit path of the tunnel.
  h. Run the **mpls te commit** command to commit the current tunnel configuration.

NOTE

Each time the MPLS TE parameters on the tunnel interface are changed, you need to run the **mpls te commit** command to commit the configuration.

i. Run the **display interface tunnel** command to query the configuration of the tunnel.

- Configure the MPLS IP tunnel.
  a. In the global config mode, run the **interface tunnel** command to create a tunnel interface and enter the tunnel interface mode.
  b. Run the **tunnel-protocol mpls ip** command to configure the tunnel protocol to MPLS IP, that is, configure the tunnel interface to work in the IP tunnel mode.
  c. Run the **source** *ip_addr* command to configure the source IP address of the tunnel. Generally, the LSR ID of the ingress is used.
  d. Run the **destination** *ip-address* command to configure the destination IP address of the tunnel. Generally, the LSR ID of the egress is used.
  e. Run the **display interface tunnel** command to query the configuration of the tunnel.

----End

12.3.2 Configuring the Tunnel Policy

Configure the tunnel selection sequence for load balancing or the tunnel binding policy in the tunnel. After the configuration is successful, packets in the tunnel are processed according to tunnel policy.

Prerequisites

The PWE3 outer tunnel must be created.

Context

The tunnel selection sequence and the tunnel binding policy are mutually exclusive. This means that you can configure only one of them.

- The IP tunnel supports the configuration of only the tunnel selection sequence.
- The MPLS TE tunnel supports the configuration of only the tunnel binding policy.

Procedure

Run the **tunnel-policy** command to create a tunnel policy name and enter the tunnel policy mode.

Step 1 For IP tunnel, run the **tunnel select-seq** command to configure the selection sequence of tunnels for load balancing.

To configure different tunnel types for load balancing according to priorities, run this command. The more the tunnel type close to keyword **select-seq**, the higher priority for load balancing.

The MA5600T/MA5603T/MA5608T does not support load balancing between different tunnels. In other words, tunnels for load balancing must be of the same type. The tunnels are selected according to the tunnel configuration.

Step 2 For MPLS TE tunnel, run the **tunnel binding** command to configure the tunnel binding policy.

To bind to a specified tunnel ID and configure the system to switch another tunnel according to the configured sequence when a tunnel is not available, run this command. After the tunnel binding policy is configured, run the **mpls te reserved-for-binding** command in the tunnel mode to allow the MPLS TE tunnel to be bound to the VPN instance.

destination *ip-addr* indicates the destination IP address of the tunnel, which must be the same as the destination IP address configured in the MPLS TE tunnel.

Step 3 In the global config mode, run the **display tunnel-policy** command to query the information about the tunnel policy.

----End

Example

To configure a tunnel policy named te_policy and bind to tunnels with the destination IP address 5.5.5.5 and IDs 10 and 20, do as follows:

```
huawei(config)#tunnel-policy te_policy
Info: New tunnel-policy is configured.
huawei(config-tunnel-policy-te_policy)#tunnel binding destination 5.5.5.5 te tunnel
 10 tunnel 20
huawei(config)#display tunnel-policy
{ <cr>|string<S><Length 1-19> }:

  Command:
         display tunnel-policy
Total   tunnel policy num:            1
Sel-Seq tunnel policy num:             0
Binding tunnel policy num:             1
Invalid tunnel policy num:             0

Tunnel Policy Name  Destination    Tunnel Intf                Down switch
------------------------------------------------------------------------------
te_policy         5.5.5.5        tunnel10                   Disable
                                 tunnel20
```

12.3.3 Configuring the PWE3 Inner PW

Configure the attribute of PW and use the PW parameters for PW binding.

Prerequisites

- MPLS L2VPN must be enabled.
- The tunnel policy must be configured.

Context

PW parameters include the following parameters: control word, jitter buffer (only for TDM PWs), maximum transmission unit (MTU), loopback IP address of the peer device, PW type, RTP control header, virtual circuit connectivity verification (VCCV), used tunnel policy, flow label classification, and TDM load time (only for TDM PWs).

Different services have different configurations when the services are bound to a PW.

Procedure

Run the **pw-para** command to create PW parameter.

PW parameters and the PW have a one-to-one mapping. One PW parameter can be used by only one PW.

Step 1 Run the **peer-address** command to configure the IP address of the peer device.

peer-address indicates the peer IP address in the PW for creating communication. In the actual transmission, data packets are automatically transmitted to the peer device according to this IP address.

Step 2 Run the **pw-type** command to configure the PW type.

The MA5600T/MA5603T/MA5608T supports TDM, ATM and ETH PWs.

The ATM PW is categorized as ATM NTo1 VCC and ATM SDU types.

- ATM NTo1 VCC: One or more ATM VCCs are transmitted on a PW.
- ATM SDU: Only the AAL5 CPCS-SDU payload is transmitted.

ETH PWs are categorized as raw and tagged modes.

- Raw mode: The PW VLAN tag is not carried in the upstream direction, but the PW payload can carry the SVLAN.
- Tagged mode: The payload of an upstream packet carries the PW VLAN tag, and the PW VLAN tag is removed in the downstream direction.

For the same PW, the PW types at both ends must be the same. In this way, the PW can be available.

NOTICE

Among PW parameters, the IP address and PW type of the peer device cannot be changed after they are configured. To change these two parameters, run the **undo pw-para** command to delete them first, and then configure them again. Make sure that the two parameters are correctly configured the first time, so as to prevent repeated operations.

Step 3 Run the **control-word** command to enable the control word mode.

When VCCV ping works in the control word mode, you need to enable the control word. It is recommended that you enable the control word mode.

Step 4 (Optional) Run the **pri-mapping-profile** command to bind an MPLS priority mapping profile to the PW.

The MPLS priority mapping profile can be configured by running the **mpls qos pri-mapping-profile** command. The profile includes the mapping from EXP to COS and the mapping from COS to EXP. To use different QoS policies based on different services for flexible mapping in the upstream and downstream directions, use this configuration.

By default, the MPLS priority mapping profile named default-profile-0 is bound to the ETH PW; the MPLS priority mapping profile named default-profile-1 is bound to the ATM PW; the MPLS priority mapping profile named default-profile-2 is bound to the TDM PW.

Step 5 (Optional) Run the **jitter-buffer** command to configure the jitter buffer.

The jitter buffer can effectively prevent jitter and delay. By default, the jitter buffer size is 2000 μs.

NOTE

- Only a TDM PW supports setting of the jitter buffer size.
- The jitter buffer size must be an integer multiple of 125.

Step 6 (Optional) Run the **mtu** command to configure the MTU.

Due to the limit in the system, the configurable MTU ranges for different PW types are different:

- MTU values set on the two devices at the ends of an ETH PW must be the same. If MTU values are different, an ETH PW can never be available.
- By default, the MTU is 1500 bytes. Do not modify this value unless there is a special requirement.

Step 7 Run the **rtp-header** command to configure the RTP control header.

NOTE

This command is applicable to only TDM PWs.

The length of the RTP header is 12 bytes, including the version number, padding flag, and timestamp fields. The timestamp field, whose length is 32 bits, is used for clock synchronization. For format of the RTP header, see RFC3550.

After RTP is enabled, PW packets of the TDM type carry the RTP control header. Otherwise, the RTP control header is not carried.

The RTP configuration must be the same as that on the peer PW device. By default, the MA5600T/MA5603T/MA5608T disables the RTP control header.

Step 8 Run the **vccv** command to enable VCCV, so as to notify the peer device of the VCCV types supported by the local device. After a successful negotiation between both devices, a virtual circuit connectivity verification is performed by using LSP ping according to the priority of the VCCV type.

VCCV is an end-to-end PW fault detection and diagnosis mechanism. Simply, VCCV is a control channel for the PW to send verification messages between the ingress and egress.

Enable the LSP ping function for alter, CW, and TTL channels or any of the three channels according to the VCCV types supported by the system. By default, VCCV is disabled.

Step 9 (Optional) Run the **tdm-load-time** command to configure the TDM load time.

NOTE

Only a TDM PW supports the setting of the load time.

Because each TDM frame is 125 μs, the load time must be an integer multiple of 125. If the entered number is not an integer multiple of 125, the system rounds it down to the nearest integer multiple of 125 μs. The jitter buffer must be greater than the load time.

The default jitter buffer is 1000 μs. Do not modify this value unless there is a special requirement.

Step 10 (Optional) Run the **tnl-policy** command to configure the tunnel policy used by the PW.

NOTE
The tunnel policy and the PW flow label classification are mutually exclusive. Configure either of them.

After the tunnel policy used by the PW is configured, the PW can perform load balancing or path selection according to the tunnel policy.

Step 11 (Optional) Run the **flow-label** command to enable flow classification.

NOTE
- The tunnel policy and the PW flow label classification are mutually exclusive. Configure either of them.
- Only the ETH PW supports flow label.
- Before configuring the flow label capability, make sure that the status of the flow label function on the local end is same as that on the peer end, and it is recommended that you adopt the same classification rules. If the flow label function is enabled on the local end but is disabled on the peer end, the packets carrying a flow label sent by the local end will be dropped after they arrive at the peer end, and the packet carrying no flow label will also be dropped after they arrive at the local end. As a result, services will be interrupted.
- After flow classification is enabled, you need to run the **mpls ecmp** command in the global config mode to enable the MPLS ECMP function. Then, the flow classification function takes effect.

To implement PWE3 load balancing, at the start point of the PW (ingress PE), the PW data is classified into different flows and each flow is allocated with a flow label. The downstream P node of the PW performs load balancing according to the flow labels.

The flow label supports the following flow classification by the source IP address, destination IP address, source MAC address, destination MAC and address, and any combination of the previous four IP addresses.

Step 12 (Optional) Run the **max-atm-cells** command to configure the maximum number of ATM cells that can be subtended.

Only the PW bound to a PW of the NTo1 VCC type requires the configuration of the maximum number of ATM cells that can be subtended. After the configuration, the number of ATM cells in the packet sent from the peer end cannot exceed this value. The default value is 1.

Step 13 (Optional) Run the **max-encapcell-delay** command to configure the packet delay of the ATM cell maximum group.

Only the PW of the NTo1 VCC type requires the configuration of the packet delay of the ATM cell maximum group. After the configuration, the maximum waiting time of subtended ATM cells encapsulated in a packet is the packet delay of the ATM cell maximum group. The default value is 0 ms.

NOTICE

If a PW is already set up and its adminstatus queried by running the **display pw** command is displayed as up, the attributes of the PW cannot be changed. Before changing the attributes, run the **manual-set pw-ac-fault** command to set the adminstatus of the PW to down. After the attributes are changed, run the **undo manual-set pw-ac-fault** command to set the adminstatus of the PW back to up. Then, the new configurations of the PW take effect.

Step 14 In the privilege mode or global config mode, run the **display pw-para** command to query the configuration of the PW.

----End

Example

To configure PW 10 with the following attributes, do as follows:

- IP address of the peer PW device: 10.10.10.20
- PW type: TDM SAToP E1
- Name of the tunnel policy used by the PW: **tdm-policy**
- Enable the RTP control header and the control word mode
- Enable the connectivity verification function of the alter, CW and TTL channels
- Other parameters: default settings

```
huawei(config)#pw-para 10
huawei(config-pw-para-10)#peer-address 10.10.10.20
huawei(config-pw-para-10)#pw-type tdm satop e1
huawei(config-pw-para-10)#tnl-policy tdm-policy
huawei(config-pw-para-10)#rtp-header
huawei(config-pw-para-10)#control-word
huawei(config-pw-para-10)#vccv cc cw alert ttl cv lsp-ping
huawei(config-pw-para-10)#quit
huawei(config)#display pw-para 10
  PW ID           : 10
  PeerIP          : 10.10.10.20
  Tnl Policy Name  : tdm-policy
  PW Type         : tdm satop e1
  CtrlWord        : enable
  VCCV Capability  : cw alert ttl/lsp-ping
  MTU            : 1500
  Statistic switch  : disable
  MaxAtmCells      : --
  MaxEncapDelay    : --
  RTP            : enable
  JitterBuffer     : 2000
  LoadTime(us)     : 1000
  TimeSlotNum      : 32
  PayLoadSize(bytes): 256
  FlowLabel Transmit          : --
  FlowLabel Classification-rule : --
  FlowLabel Receive           : --
  Priority mapping profile name : default-profile-2
```

To configure PW 20 with the following attributes, do as follows:

- IP address of the peer PW device: 10.20.30.40
- PW type: ETH Tagged
- Name of the tunnel policy used by the PW: **eth-policy**
- Other parameters: default settings

```
huawei(config)#pw-para 20
huawei(config-pw-para-20)#peer-address 10.20.30.40
huawei(config-pw-para-20)#pw-type ethernet tagged
```

```
huawei(config-pw-para-20)#tnl-policy eth-policy
huawei(config-pw-para-20)#quit
huawei(config)#display pw-para 20
  PW ID           : 20
  PeerIP          : 10.20.30.40
  Tnl Policy Name   : eth-policy
  PW Type         : ethernet tagged
  CtrlWord        : disable
  VCCV Capability   : disable
  MTU             : 1500
  Statistic switch  : disable
  MaxAtmCells      : --
  MaxEncapDelay    : --
  RTP             : --
  JitterBuffer     : --
  LoadTime(us)     : --
  TimeSlotNum      : --
  PayLoadSize(bytes): --
  FlowLabel transmit          : disable
  FlowLabel classification-rule : --
  FlowLabel receive           : disable
  Priority mapping profile name : default-profile-0
```

12.3.4 Binding the Service to the PW

Bind various PWE3 services to a PW. After the binding, user packets are encapsulated and forwarded according to the modes defined in the PW parameters.

Prerequisites

- The PW must be configured.
- For TDM PWE3, the TDM connection must be created.
- For ATM PWE3, the ATM-based service port must be created.
- For ETH PWE3, the ETH-based service port must be created.

Context

Different PWE3 services have different configurations when the services are bound to a PW.

- TDM PWE3 supports dynamic PW and static PW.
- ATM PWE3 supports dynamic PW and static PW.
- ETH PWE3 supports dynamic PW and static PW.

The parameters of a static PW are not negotiated using the signaling protocol, the relevant information is configured manfully through the command line interface (CLI), and the data is transmitted through tunnels between PEs.

Procedure

- Bind the TDM service to a PW.

  Run the **pw-ac-binding tdm** command to use a PW to create the TDM PW service.

  Pay attention to the following points during the configuration:

  - To specify a PW as a static PW, you need to configure the in label and out label of the PW. The out label value must be an unallocated and idle value at the peer end and the in label value must be an unallocated value at the local end.
  - To specify a PW and an UDP PW, you need to configure the destination port ID and source port ID of the PW. The destination port ID must be the same as the source port ID at the peer PW device and the source port ID must be the same as the destination port ID at the peer PW device.

- Bind the ATM service to a PW.

  Run the **pw-ac-binding pvc** command to use a PW to create the ATM PW service.

  The PVC and the PW can be bound in two modes: NTo1 mode and SDU mode. Pay attention to the following points during the configuration:

  - In the SDU mode, a PW is bound to only one PVC. Therefore, you need not change the VPI or VCI.
  - In the NTo1 mode, a PW can be bound to multiple PVCs. To differentiate between PVCs, you must change the out VPI and VCI of the PW, that is, you must specify **outvpi** and **outvci**. Operation procedure is as follows:
    i. Run the **pw-ac-binding pvc** command to bind a PW to a PVC.
    ii. Run the **pw-ac-append pvc** command to bind the PW to another PVC.

- Bind the ETH service to a PW.

  Run the **pw-ac-binding vlan** command to use a PW to create the ETH PW service.

  Note: To specify a PW as a static PW, you need to configure the in label and out label of the PW. The out label value must be an unallocated and idle value at the peer end and the in label value must be an unallocated value at the local end.

----End

Example

To create a static binding between TDM connection 0 and PW 20 (outgoing label/incoming label: 16/8448), do as follows:

```
huawei(config)#pw-ac-binding tdm 0 pw 20 static transmit-label 16 receive-label 8448
```

To bind the ATM service to a PW with the following settings, PW type to ATM sdu, do as follows. Settings: ATM access port 0/3/0, VPI/VCI 0/35, and PW ID 20.

```
huawei(config)#pw-ac-binding pvc 0/3/0 vpi 0 vci 35 pw 20
```

To bind the ETH service to a PW with the following settings, do as follows. Settings: VLAN ID 100, PW ID 30, PW out label 8500, and PW in label 8600.

```
huawei(config)#pw-ac-binding vlan 100 pw 30 static transmit-label 8500 receive-label
8600
```

12.3.5 Configuring PW Protection

Create a standby PW for a PW. When the active PW is faulty, the system quickly switches to the standby PW to ensure the service reliability.

Prerequisites

- The active PW must be created.

- The basic parameters are configured. For the configuration method, see 12.3.3 Configuring the PWE3 Inner PW.

Context

PW protection: When a PW is faulty (such as an LDP session is down, a tunnel is deleted, the protocol communication is faulty, the route status changes, or VCCV has no response), the system can quickly switch to the standby PW. Then, the standby PW functions as the active PW.

The MA5600T/MA5603T/MA5608T supports PW 1:1 redundancy.

Procedure

Run the **pw-protect** command to configure the standby PW.

Pay attention to the following points during the configuration:

- The standby PW ID cannot exist.
- The PW parameters of the active and standby PWs must be the same.
- Both active PW and the standby PW are not static PW.

Step 1 (Optional) Enable the PW protection group to support dual-sending and dual-receiving for the multicast service.

When the PW protection group supports dual-sending, both active and standby PWs can forward IGMP packets so that the multicast forwarding entry can also be created on the device corresponding to the standby PW. After the active/standby PW switchover, the multicast service can be smoothly switched. This configuration is recommended when the multicast service is carried by the active and standby PWs.

When the PW protection group supports dual-receiving, both active and standby PWs can receive packets to avoid packet loss caused by signaling delay when switchback is performed after the faulty active PW recovers. This configuration is recommended when the multicast service is carried by the active and standby PWs.

1. Run the **igmp_send_dual-pw** command to set whether IGMP packets can be sent by both active and standby PWs.
2. Run the **pw-redundancy_stream-dual-receiving** command to set the PW protection group to work in the dual-receiving mode.

Step 2 Run the **pw-revertive-mode** command to configure the switchback policy for the PW protection group.

Switchback: When both active and standby PWs are available, if the original service traffic is carried on the standby PW, the service can be switched back to the active PW according to actual requirements. Set the switchback policy according to actual network conditions (such as whether the network topology often changes and whether the traffic should be carried on the active PW).

The switchback policy of a PW protection group can be immediate automatic switchback, automatic switchback after a period of time, and no automatic switchback.

Step 3 Run the **display pw-ps** command to query the configuration of the PW protection group.

----End

Example

To configure a PW protection group, set the parameters as follows: active PW ID to 10, standby PW ID to 20, and switchback policy to allowing automatic switchback for the PW protection group in 30 seconds.

```
huawei(config)#pw-protect primary-pw 10 secondary-pw 20
huawei(config)#pw-revertive-mode 10 revertive wtr 30
huawei(config)#display pw-ps 10
  ---------------------------------------------------------------------------
  Primary-PW-ID    Primary-PW-state    Secondary-PW-ID    Secondary-PW-state
  ---------------------------------------------------------------------------
            10          up/active               20                down
  ---------------------------------------------------------------------------
  revertive-mode: revertive, in 30 seconds
```

12.3.6 Configuring MPLS Tunnel Protection

Create a protection tunnel for the MPLS TE tunnel. When the working tunnel is faulty, the system quickly switches to the protection tunnel to ensure the service reliability.

Prerequisites

- The forward LSP must be created.
- The backward LSP must be created.
- MPLS OAM must be enabled.

Context

MPLS tunnel protection is a part of the MPLS OAM connectivity detection mechanism.

The basic process of the MPLS OAM connectivity check and protection switching is as follows:

1. The source transmits the CV/FFD packets to the destination through the detected LSP.
2. The destination checks the correctness of the type and frequency carried in the received detection packets and measures the number of correct and errored packets that are received within the detection period to monitor the connectivity of the LSP in real time.
3. After detecting a defect, the destination transmits the BDI packets that carry the defect information to the source through the backward path.
4. The source learns about the status of the defect, and triggers the corresponding protection switching when the protect group is correctly configured.

Procedure

Configure working MPLS TE tunnel.

1. In global config mode, run the **interface tunnel** command to create a tunnel interface and enter the tunnel interface mode.
2. Run the **tunnel-protocol mpls te** command to configure the tunnel protocol to MPLS TE.
3. Run the **destination** *ip-address* command to configure the destination IP address of the tunnel. Generally, the egress LSR ID is used.

4. Run the **mpls te tunnel-id** command to configure the tunnel ID.
5. Run the **mpls te signal-protocol rsvp-te** command to configure the signaling protocol of the tunnel to RSVP-TE.
6. (Optional) Run the **mpls te bandwidth** command to configure the bandwidth for the tunnel. After the configuration is completed, only the VLAN interface that meets this bandwidth value can be selected as the node traversed by the MPLS TE tunnel path when the MPLS TE tunnel is created.

   If the MPLS TE tunnel is only used to change the data transmission path, you may not configure the tunnel bandwidth.
7. (Optional) Run the **mpls te path explicit-path** command to configure the explicit path used by the MPLS TE tunnel.

   If only the bandwidth used by the MPLS TE tunnel is limited but the transmission path is not limited, you may not configure the explicit path used by the MPLS TE tunnel.
8. Run the **mpls te commit** command to commit the current configuration of the tunnel.

Step 1 Configure protection MPLS TE tunnel.

The working mode of MPLS OAM protection switching is 1:1 protection. Normally, each working tunnel has a protection tunnel.

The configuration of the protection tunnel is the same as that of the working tunnel.

Step 2 Configure a tunnel protect group.

Configure the working tunnel and the protection tunnel as a tunnel protect group. When the source end finds the active LSP is defective through the MPLS OAM detection mechanism, and the protection switching is required, the system can switch the data to the protection tunnel for continuous transmission.

1. In the global config mode, run the **interface tunnel** command to enter the working tunnel interface mode.
2. Run the **mpls te protection tunnel** command to create a tunnel protect group and set the switchback mode of the protect group.

   The switchback policy of a PW protect group can be immediate automatic switchback, automatic switchback after a period of time, and no automatic switchback.

Step 3 (Optional) Run the **mpls te protect-switch** command forcibly switch over the tunnel protect group.

To manually switch data streams between working and protection tunnels, run this command.

There are for forcible switching modes:

- **clear**: clears all external switching commands that are already executed in the system.
- **lock**: lock switching, which locks data streams on the working tunnel.
- **force**: forcible switching, which forcibly switch data streams to the protect tunnel.
- **manual work-lsp**: manually switches data streams on the working tunnel to the protection tunnel.
- **manual protect-lsp**: manually switches data streams on the protection tunnel to the working tunnel.

Keywords **clear**, **lock**, **force**, and **manual** corresponds to switching priorities in descending order. If a command with a higher priority is executed, a command with a lower priority cannot be executed.

Step 4 In the global config mode, run the **display mpls te protection tunnel** command to query the configuration of the tunnel protect group.

----End

Example

To configure RSVP-TE tunnel IDs to 10 and 30, destination IP address of the tunnels to 3.3.3.3, tunnel 30 as the protection tunnel of tunnel 10, switchback mode to revertive, and WTR time to 900s, do as follows:

```
huawei(config)#interface tunnel 10
huawei(config-if-tunnel10)#tunnel-protocol mpls te
huawei(config-if-tunnel10)#destination 3.3.3.3
huawei(config-if-tunnel10)#mpls te tunnel-id 10
huawei(config-if-tunnel10)#mpls te signal-protocol rsvp-te
huawei(config-if-tunnel10)#mpls te bandwidth ct0 5120  //(Optional) Configure the
global bandwidth of tunnel 10 to 5210 kbit/s.
huawei(config-if-tunnel10)#mpls te commit
huawei(config-if-tunnel10)#quit

huawei(config)#interface tunnel 30
huawei(config-if-tunnel30)#tunnel-protocol mpls te
huawei(config-if-tunnel30)#destination 3.3.3.3
huawei(config-if-tunnel30)#mpls te tunnel-id 30
huawei(config-if-tunnel30)#mpls te signal-protocol rsvp-te
huawei(config-if-tunnel30)#mpls te bandwidth ct0 5120  //(Optional) Configure the
global bandwidth of tunnel 30 to 5210 kbit/s.
huawei(config-if-tunnel30)#mpls te commit
huawei(config-if-tunnel30)#quit

huawei(config)#interface tunnel 10
huawei(config-if-tunnel10)#mpls te protection tunnel 30 mode revertive wtr 30
huawei(config-if-tunnel10)#mpls te commit
huawei(config-if-tunnel10)#quit
```

12.3.7 Configuring PW-based trTCM by CoS Remarking

In an asynchronous transfer mode (ATM) multi-protocol label switch (MPLS) network, quality of service (QoS) is required for the user ATM cells carried in ATM pseudo wire emulation edge to edge (PWE3) over the packet switched network (PSN) network. Due to mechanism differences, two rate three color marker (trTCM) by class of service (CoS) remarking is implemented on the ingress PE (MA5600T/MA5603T/MA5608T) to map ATM traffic policing mechanism to the MPLS traffic policing mechanism, and CoS-based early drop is implemented on the egress PE.

Prerequisite

The specified PW is configured.

Context

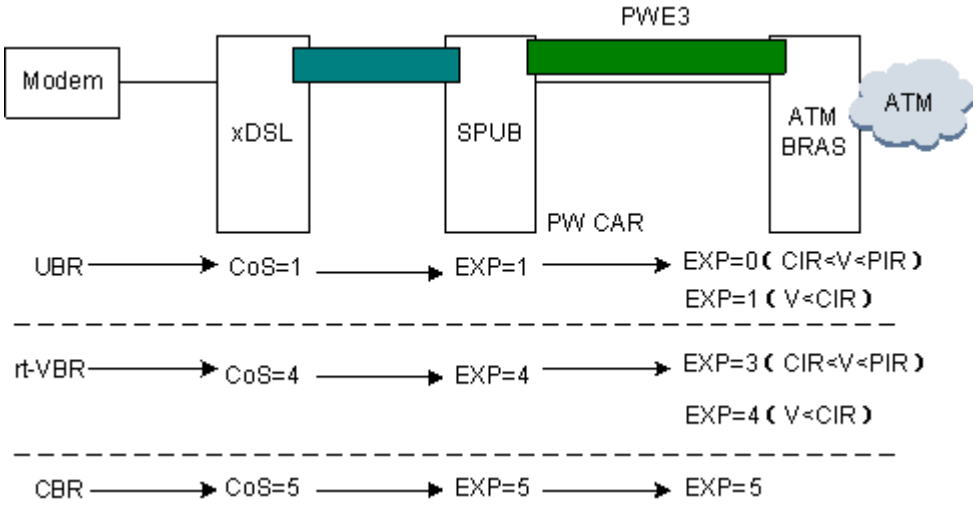
In the ATM MPLS network, certain upper-layer ATM devices fail to identify the DEI bit in the VLAN Tag field and therefore fail to implement trTCM by the DEI bit for ATM private line user. In this case, in the upstream direction of an ingress PE

(MA5600T/MA5603T/MA5608T), trTCM by CoS remarking is performed according to PW committed access rate (CAR). With this mechanism, MPLS packets whose rate is lower than committed information rate (CIR) are green packets. They are marked with the default CoS value, and MPLS packets who rate is higher than CIR and lower than peak information rate (PIR) are yellow packets. Their CoS values are remarked by the system by running the **cos-remark** command. while MPLS packets who rate are red packets. They are directly discarded by the system. When a PW packet is being encapsulated, the CoS of the MPLS packet is mapped to the EXP field of the outer MPLS label (the EXP field of the inner PW label is also processed in this way). Then, traffic management is performed in the PSN network based on the EXP field of the MPLS label.

In the downstream direction of an egress PE, if the color policy of the traffic profile is **cos**, packets are discarded based on CoS values in the EXP field. The packets carrying the default CoS value and remarking low-priority CoS values are mapped to the same queue. Then, different early-drop thresholds are configured for packets with different colors in the queue to ensure that packets whose rate is lower than CIR have a higher priority when congestion occurs.

In the upstream direction, the MA5600T/MA5603T/MA5608T implements PW-based trTCM for CIR and PIR by CoS remarking on the SPUB board, as shown in Figure 12-46. In the downstream direction, the MA5600T/MA5603T/MA5608T does not perform CAR or CoS-based early drop on the SPUB board, but implements queue-based early drop on the xDSL board according to the CoS early drop threshold.

Figure 12-46 PW-based trTCM on the upstream SPUB board



NOTE

- V indicates the rate of the MPLS packet.
- In the downstream direction, the CoS value processing on theMA5600T/MA5603T/MA5608T is reverse to that in the upstream direction.

Procedure

- Configure trTCM by CoS remarking in the upstream direction.
  a. Run the **mpls car-pw** command to configure PW-based rate limitation.

In the upstream direction of an ingress PE, the system remarks packets with different rates with different colors based on the CIR and PIR parameters and discards the red packets.

- Only the upstream packets on the SPUB board are remarked.
- The CAR parameters of a PW can be configured only after the MPLS L2 VPN function is enabled by running the **mpls l2vpn** command.
- CIR is mandatory, and the other three parameters are optional. If you configure only the CIR, the system calculates the other three parameters based on the formula. It is recommended to configure only the CIR. The relationships between these parameters are as follows:
  - CBS = (CIR+8191)/8192+1
  - PIR = 2*CIR
  - PBS = MAX(((PIR+8191)/8192+1), CBS)

b. Run the **cos-remark** command to configure the priority remarking policy of trTCM.

The system remarks the priorities of yellow packets with a low-priority CoS value. In the downstream direction of an egress PE, if the color policy of the traffic profile used by the traffic stream is **cos**, priority-based early drop is implemented.

- By default, the CoS value remarked is the same as the original value.
- The CoS value remarked is low-priority CoS value while the other CoS value is the high-priority CoS value. The high-priority CoS value can be remarked as the low-priority CoS value while the low-priority CoS value can only be remarked as itself. For example, when priority A is remarked as priority B, A is the high-priority CoS value and B is the low-priority CoS value. Priority B can only be remarked as priority B.

- Configure CoS-based early drop in the downstream direction.

a. Run the **traffic table ip** command to create an IP traffic profile for AoE streams. To implement CoS-based early drop for AoE streams, the color policy must be set to **cos**.

- Only the ADSL and SHDSL boards support early drop for the downstream packets.
- The default color policy is **dei**.
- The color policy in the upstream and downstream traffic profiles used by the traffic stream must be the same.
- **traffic table ip** indicates traffic stream-based rate limitation and **mpls car-pw** indicates PW-based rate limitation. If more than one rate limitation modes are configured in the system, the minimum rate is used.
- CIR is mandatory, and the other three parameters are optional. If you configure only the CIR, the system calculates the other three parameters based on the formula. It is recommended to configure only the CIR. The relationships between these parameters are as follows:
  - CBS = min(2000+CIR*32, 10240000)
  - PIR = min(2*CIR, 10240000)
  - PBS = min(2000+32*PIR, 10240000)

b. Run the **cos-queue-map** command to configure the mapping between queues and CoS values (802.1p priority) so that packets with different priorities are mapped to the specified queues based on the configured mapping. This enhances the flexibility of mapping packets to the queue.

- The larger the queue ID, the higher the priority for forwarding packets.
- The default mapping between priorities and queue IDs is: Priority 0 maps to queue 0; priority 1 maps to queue 1; the same rule applies to other priorities.
- A PVC may have two different CoS values: a default CoS value and a remarked CoS value. To ensure that packets with these two CoS values are in correct sequence during AoE encapsulation, the packets are mapped to the same queue.
- It is recommended that the same type of permanent virtual paths (PVCs) be encapsulated into the same PW. Otherwise, different types of PVCs have the same EXP value and service-based queue scheduling cannot be implemented.

c. Run the **wred-profile** command to add a weighted random early detection (WRED) profile. Configure the early-drop thresholds and drop ratios for the green and yellow packets.

d. Run the **queue-wred** command to bind the WRED profile to queues. After the WRED profile is bound, the system performs color-based early drop according to the parameters configured in the WRED profile.

----End

Example

Assume that in the upstream direction of the MA5600T/MA5603T/MA5608T, CIR of PW 10 is 1 Mbit/s and yellow packets are remarked with CoS 0; in the down stream direction, green packets are not dropped, low drop threshold for yellow packets is 50, high drop threshold 80, and drop ratio 100. In an ATM MPLS network, to create the AoE service port (no CAR for the service port) by referencing traffic profile 8 (in which the VLAN priority is set to 1) to provision 1 Mbit/s service the ATM private line user, do as follows:

```
huawei(config)#mpls car-pw 10 cir 1024
huawei(config)#cos-queue-map cos1 0 cos4 3
huawei(config)#traffic table ip index 8 name "CBR" cir off color-policy cos priority
1 priority-policy tag-in-package
huawei(config)#service-port 2 vlan aoe adsl 0/2/0 vpi 0 vci 35 single-service inbound
traffic-table index 8 outbound traffic-table index 8
huawei(config)#cos-queue-map cos0 0 cos1 0 cos3 1 cos4 1 cos5 5 cos2 3 cos6 3 cos7 3
huawei(config)#wred-profile index 6 green low-limit 100 high-limit 100
discard-probability 0 yellow low-limit 50 high-limit 80 discard-probability 100
huawei(config)#queue-wred queue0 6 queue1 6 queue2 6 queue3 6
```

nx

12.3.8 Configuring CR-LSP Backup

Backup CR-LSPs are established on networks requiring high reliability to provide end-to-end protection, ensuring network reliability. If a primary CR-LSP fails, traffic rapidly switches to a backup CR-LSP, ensuring uninterrupted traffic transmission.

Prerequisites

Before configuring CR-LSP backup, complete the following tasks:

- 10.7.3 Configure an RSVP-TE LSP

- Enabling MPLS, MPLS TE, and RSVP-TE globally and on interfaces of each node along a backup CR-LSP

Context

Hot standby and ordinary backup modes are supported. If both primary and backup CR-LSPs fail, best-effort paths are established. The following table lists CR-LSP backup modes.

Backup Mode	Description	Advantage	Shortcoming
Hot standby	A hot-standby CR-LSP is set up over a separate path immediately after a primary CR-LSP is set up	A rapid traffic switchover can be performed	Additional bandwidth needs to be reserved for a hot-standby CR-LSP
Ordinary backup	The system attempts to set up an ordinary backup CR-LSP if a primary CR-LSP fails	No additional bandwidth is needed	Ordinary backup performs a traffic switchover slower than hot standby
Best-effort path	The system establishes a best-effort path over an available path if both the primary and backup CR-LSPs fail	Establishing a best-effort path is easy and a few constraints are needed	Some quality of service (QoS) requirements cannot be met

Tthe following backup modes are supported for CR-LSP.

- Hot standby
- Ordinary backup
- Hot standby + best-effort path

Procedure

Configure the backup mode of CR-LSP.

1. Run the **interface tunnel** command to enter MPLS TE tunnel mode.
2. Run the **mpls te backup** command to configure the backup mode of current tunnel.
   - The keyword **hot-standby** indicates the hot standby mode.
   - **ordinary** indicates the ordinary mode.
   - **ordinary best-effort** indicates the best-effort path mode.

Step 1 (Optional) Configure the explicit path for a backup CR-LSP.

After hot standby or ordinary backup is configured, the system automatically selects a path for a backup CR-LSP. To manually specify a path for a backup CR-LSP, you can set explicit path. An explicit path consists of a series of nodes, which constitute a vector path according to the configured sequence. The IP address in an explicit path is the IP address of the interface on the node. Generally, the loopback interface IP address on the egress is used as the destination IP address of the explicit path.

NOTE

Use a separate explicit path for the backup CR-LSP to prevent the backup CR-LSP from completely overlapping its primary CR-LSP. Protection will fail if the backup CR-LSP completely overlaps its primary CR-LSP.

1. Run the **explicit-path** command in the global config mode to create an explicit path.
2. Run the **next hop**, **modify hop**, and **delete hop** command to add a next hop node, modify a node, and delete a node respectively for the explicit path.
3. In the tunnel mode, run the **mpls te path explicit-path** *path-name* **secondary** command to specify the explicit path for the backup CR-LSP.

Step 2 (Optional) Configure the affinity property for a backup CR-LSP.

Affinity property masks determine the link properties that should be checked by a device. To ensure that a link can be used by a tunnel, for the bits that are 1 in a mask, it is required that at least one bit in the administrative group and the corresponding bit in the affinity property be 1. In addition, if the bits in the affinity property are 0, the corresponding bits in the administrative group cannot be 1.

The default affinity property is 0x0.

Run the **mpls te affinity property secondary** command to set the affinity property for the backup CR-LSP.

Step 3 (Optional) Configure the hop limit for a backup CR-LSP.

After a CR-LSP is configured with a hop limit, the hop limit acts as one of routing conditions such as the link bandwidth and affinity property when the CR-LSP is created. After the hop limit is set, the number of hops of a CR-LSP cannot exceed this limit.

The default hop limit is 32.

Run the **mpls te hop-limit secondary** command to set hop limit for the backup CR-LSP.

Step 4 Run the **mpls te commit** command to save configuration.

----End

Follow-up Procedure

After the configuration of CR-LSP backup is finished, you can query information about the tunnel interface and backup status.

- Run the **display mpls te tunnel-interface** command to check information about a tunnel interface.
- Run the **display mpls te hot-standby state** command to check information about the hot-standby status.
- Run the **display mpls te tunnel** command to check CR-LSP information.

13 Layer 2 Forwarding

13.1 Overview

The Layer 2 forwarding feature includes the following sub features: MAC address management, VLAN, service flow, service port bundle, forwarding policy, and access user bridging.

Layer 2 Forwarding Solutions

Table 13-1 Layer 2 Forwarding Solutions

Feature	Description	Application Scenario
13.2 MAC Address Management	MAC address management is a basic Layer 2 management feature, including setting the MAC address aging time, limiting the number of dynamic MAC addresses (the number of the MAC addresses that can be learned), and setting the static MAC address.	• The system ages dynamic MAC addresses to ensure timely updates of the MAC address table. If the MAC address table is full and not updated, the system will fail to learn new MAC addresses and will consequently fail to forward data. • By limiting the number of learnable dynamic MAC addresses, the system administrator can limit the number of MAC addresses that enter the network and hence alleviate the load of network devices. • By configuring static MAC addresses, the system administrator denies access to unauthenticated users.
13.3 VLAN	VLAN allows packets to be broadcast only within a single VLAN, preventing bandwidth waste caused by broadcast storms. Furthermore, VLAN enhances network security, because NEs in different	For example, for users of different enterprises in the same office building, it is too costly to build separate LANs for each enterprise, and insecure if the enterprises are to share the existing LAN of the building. VLAN resolves this dilemma. Make different enterprise users

Feature	Description	Application Scenario
	VLANs cannot communicate with each other.	belonging to different VLANs. In this way, the enterprise users share the LAN facilities and at the same time each have their own secure networks.
13.4 Service Flow	Service flow, also called service port, is a result of traffic classification by characteristics of an Ethernet packet on a physical or logical port. Service flow is also a Layer 2 logical channel that carries services between the MA5600T/MA5603T/MA5608T and users (Specify the Layer 2 forwarding path).	An access device provides services to a large number of users, and each user requires multiple types of services (for example, HSI and VoIP services). The access device is required to differentiate between different user services when processing user packets, so that the services do not interfere with each other. To address this requirement, the MA5600T/MA5603T/MA5608T provides the service flow feature. In addition to traffic classification, a service flow is the smallest unit of user service processing. Hence, differentiated and fine-grained management, such as QoS, line identification, and security policies, of user services can be implemented based on service flows.
13.5 Service Port Bundle	Service port bundle, also named service flow bundle or flow bundle, is a CoS-based packet forwarding model. Each service port bundle corresponds to a group of services for a user. Each service flow carries one type of service and has a CoS level. A service port bundle can also be considered a bundle of service flows. The following figure shows the schematic diagram for service port bundle.	In a network where VLANs are planned based on ports on the access node and the access node connects to users through a router, the packets transmitted from a user to the access node through the router may carry the same MAC address even when they carry different services. As a result, MAC address transfer may occur between different service flows (service flow 1 and service flow 2) on the access node. The service port bundle feature addresses this issue using CoS-based route selection.
13.6 Layer 2 Forwarding Policy	As a Layer 2 network device, the MA5600T/MA5603T/MA5608T forwards packets based on S-VLAN+C-VLAN on Layer 2 networks. In this packet forwarding mode, packets are forwarded based on VLANs and Layer 2 forwarding mapping is based on S-VLAN+C-VLAN IDs but not MAC address learning.	In traditional Layer 2 packet forwarding, packets are forwarded based on their VLAN and MAC address (VLAN+MAC address). If the destination MAC address of the packets become invalid because of dynamic MAC address aging, the MA5600T/MA5603T/MA5608T fails to query VLAN+MAC address. As a result, the MA5600T/MA5603T/MA5608T broadcasts the packets as unknown unicast packets in the VLAN, which brings in security threats. In addition, the Layer 2 packet forwarding based on

Feature	Description	Application Scenario
		VLAN+MAC address faces security issues caused by MAC address spoofing and attacks. The Layer 2 packet forwarding based on S-VLAN+C-VLAN can resolve the preceding issues.
13.7 Layer 2 User Bridging	After Layer 2 bridging is enabled, all users connected to an MA5600T/MA5603T/MA5608T can exchange data on Layer 2 networks.	The QinQ service deployment requires that users can communicate with each other on Layer 2 networks. However, all users connected to an MA5600T/MA5603T/MA5608T are isolated on Layer 2 networks. Due to this limitation, the QinQ service can be deployed only between MA5600T/MA5603T/MA5608Ts. Users connected to an MA5600T/MA5603T/MA5608T and in different network segments require an upper-layer device to forward data at Layer 3 to exchange data. This requires that the upper-layer device support ARP proxy, promoting higher requirements on the upper-layer device. The Layer 2 bridging can resolve the preceding issues.

13.2 MAC Address Management

MAC address management is a basic Layer 2 management feature, including setting the MAC address aging time, limiting the number of dynamic MAC addresses that can be learned, and setting the static MAC address.

13.2.1 What Is MAC Address Management

Definition

MAC address management is a basic Layer 2 management feature that enables system administrators to use the functions listed in the following table.

Table 13-2 Sub-functions of MAC address management

Sub-function of MAC Address Management	Description	Remarks

Sub-function of MAC Address Management	Description	Remarks
Setting the MAC address aging time	After a system administrator sets the MAC address aging time, the system periodically checks for aged dynamic MAC addresses. If the system detects no packets, whether sent or received, carrying specified source MAC addresses within one or two times of the aging time, the system deletes the MAC address from the MAC address table.	The system saves the MAC address table in its buffer. MAC address entries specify the mapping between the MAC addresses of devices, port numbers, and VLAN IDs. When forwarding frames, the system consults the MAC address table according to the destination MAC addresses and VLAN IDs of these frames and quickly identifies the egress for them. This function prevents frame broadcast. MAC addresses in a MAC address table can be either manually configured or dynamically learned by the system.
Limiting the number of learnable dynamic MAC addresses	The system allows the system administrator to configure the number of learnable dynamic MAC addresses on a port or a service flow. When the number of learned MAC addresses reaches the maximum number, the port or service flow does not learn any new MAC addresses.	The system can age out dynamic MAC addresses. When the system does not receive packets from or send packets to a device within a specified period, the system deletes the MAC address entry of the device from the MAC address table.
Setting the static MAC address	To connect a port to a device with a specified MAC address, the system administrator configures a static MAC address and VLAN on a port or configure a static MAC address on a service port. The system then forwards data according to this static MAC address.	None

Benefits

Benefits for Carriers

- The system ages dynamic MAC addresses to ensure timely updates of the MAC address table. If the MAC address table is full and not updated, the system will fail to learn new MAC addresses and will consequently fail to forward data.
- By limiting the number of learnable dynamic MAC addresses, the system administrator can limit the number of MAC addresses that can be used to enter the network and hence alleviate the load of network devices.

- By configuring static MAC addresses, the system administrator prohibits unauthorized users from accessing the system.

NOTE

- If a malicious user sets source MAC addresses of attack packets to the MAC addresses of authorized users and accesses the system through another port on the device, the device learns an incorrect MAC address entry and forwards to the malicious user the packets that were originally destined to an authorized user. To avoid this situation, the system administrator can manually add specific MAC address entries to the MAC address table to bind user devices to ports. This operation protects user data from MAC spoofing and enhances port security. The manually added MAC address entries (static MAC address entries) have higher priorities than automatically generated MAC address entries.
- The static MAC addresses, however, do not automatically update according to network device changes. The system administrator needs to modify the static MAC addresses manually. Using a large number of MAC addresses will require higher costs for network maintenance.

Benefits for Subscribers

Improved user security: After the system administrator sets the static MAC address of a service port and sets the maximum number of learnable MAC addresses to 0, the port receives only user data carrying the specified static MAC address.

13.2.2 MAC Address Management Process

MAC address management includes MAC address table establishment and management

Establishing MAC Address Tables

The system establishes a MAC address table by learning source MAC addresses or after users configure static MAC address entries.

- **MAC address learning**

  When functioning as a Layer 2 switch, the MA5600T/MA5603T/MA5608T learns the source MAC address of the packets received by its ports and forwards these packets according to their destination MAC addresses. The device saves the learned MAC addresses in its buffer. Generally, buffer can hold a limited number of MAC addresses. If all the buffer is full, no more MAC addresses can be learned.

NOTE

The MAC address learning function takes effect only for unicast MAC addresses, but not for multicast or broadcast MAC addresses.

The MAC address learning function can be configured for the control board and service boards separately.

Table 13-3 The classification of the MAC address learning function

Classification	Description	Available scene	Configuration Command
MAC address learning on the control board	The control board learns the MAC addresses of the transmitted and received packets.	• If the control board forwards packets in S+C mode and one S-VLAN includes multiple service boards, it is recommended that you enable MAC address learning on the control board. This function helps avoid the duplication and broadcast of excessive unknown	• **mac-address learning fabric** • **mac-address learning all**

Classification	Description	Available scene	Configuration Command
		unicast packets. • In cascading scenarios, if the control board has insufficient space for MAC addresses, it is recommended that you disable MAC address learning on the control board.	
MAC address learning on the service board	A service board whose MAC address function is enabled learns the source MAC addresses and VLAN IDs of the received packets.	• MAC address learning is preferred when the system forwards packets by VLAN+MAC. • If excessive MAC addresses are learned from a VLAN and the service board has insufficient space for holding more MAC addresses, it is recommended that you disable the VLAN level MAC address learning function on the service board.	**mac-address learning vlan**

- **Configuration of static MAC address entries**
  - A user can manually configure static MAC address entries in which user device MAC addresses are bound to ports. After this configuration, the packets whose MAC addresses are included in the MAC address entries are always forwarded through the bound ports. This configuration improves the efficiency for forwarding packets and improves the security of ports because it denies access from unauthenticated users. This method of establishing MAC address tables is widely used in private networks.
  - The same static MAC address can be configured on an upstream port belonging to different VLANs.
  - Configuration command: **mac-address static**

The following table shows an example of a simplified MAC address table established by configuring static MAC address entries. The table lists the mapping between MAC addresses, ports, and VLAN IDs.

Table 13-4 Simplified MAC address table

VLAN ID	MAC Address	Port Number	MAC Address Attribute	Forwarding Attribute
2	0000-0010-0011	0/1/1	Dynamic	Forward
3	0000-0020-0033	0/1/2	Static	Forbid

Managing MAC Address Tables

When managing MAC address tables, users can configure MAC-related attributes as allowed by system resources and network security policies against potential risks. The optimized MAC address tables can better meet requirements of a live network. These MAC-related attributes are as follows:

- **Maximum number of MAC addresses learned based on service flows**
  - After the number of access users reaches the limit, no new access user addresses will be learned. This attribute setting applies to networks, such as residential access networks and low-security internal enterprise networks, that have fixed access users but are not sufficiently secure.

NOTE

- The maximum number of MAC addresses is independent of the number of manually added static MAC addresses.
- After a static MAC address is configured for a user port and the maximum number of learnable dynamic MAC addresses is set to 0, the port receives only packets that contain the configured static MAC address. In this case, the MAC address is bound to the port. This attribute setting helps improve user data security.

  - Specifically, after a service flow is added to a bundle, the system uses the maximum number for the bundle as that for the service flow. The maximum number of learnable MAC addresses for the entire bundle is the same as that for each service flow in the bundle. When the maximum number for any service flow in a bundle changes, the maximum number for each bundle and that for each service flow in the bundle change as well. The sum of MAC addresses learned for each service flow in a bundle cannot exceed the maximum number of learnable MAC addresses for the bundle.
  - Configuration command: **mac-address max-mac-count**

- **Setting the rate for learning MAC addresses**

  A lot of CPU resources will be used if the system learns MAC addresses at a high rate, affecting other services. If the system learns MAC addresses at a low rate, it may fail to learn MAC addresses, causing excessive unknown unicasts. To resolve the issues, a proper rate for learning MAC addresses must be specified.

  After the rate for learning MAC addresses is specified, the system forwards the packets that exceed this limit and does not learn their MAC addresses.

  Configuration command: **mac-address learning-rate**

- **Setting the function of sensing excess MAC addresses**

  When a lot of MAC addresses are learnt by the system, it is difficult for trouble locating. When the function of sensing excess MAC addresses is enabled, the system queries MAC addresses every 15 minutes and determines whether an alarm needs to be reported according to the query result. If the query result exceeds the upper threshold for sensing excess MAC addresses set by users, an excess MAC address alarm is generated. If the query result is smaller than the lower threshold for sensing excess MAC addresses set by users, a fault clearing alarm is generated.

  Configuration command: **overload-aware mac-address**

- **MAC address aging**

  Generally, the system automatically establishes a MAC address table by learning source MAC addresses. The established MAC address table has to be updated according to network changes. However, after the network topology changes, the dynamic MAC address entries will not be automatically updated in a timely manner. Then the system cannot learn more MAC addresses and user data cannot be forwarded as expected. MAC address aging is intended to resolve the issue.

NOTE
Disable MAC address aging if no MAC address aging issue has occurred.

Item	Description	Configuration Command
Specify the aging time of a MAC address	After a system administrator sets the MAC address aging time, the system periodically checks for aged dynamic MAC addresses. If the system detects no packets, whether sent or received, carrying specified source MAC addresses within one or two times of the aging time, the system deletes the MAC address from the MAC address table. Then the system creates another MAC address table by learning MAC addresses.	**mac-address timer**
Sets the aging mode for a MAC address	After you set the MAC address aging mode, the control board and service board age MAC addresses accordingly. • Bidirectional aging mode: The system ages MAC address entries according to the source MAC address learning and receiving of packets destined for this MAC address. Specifically, if the system neither learns a source MAC address nor receives packets destined for this MAC address within an aging period, the system ages the MAC address entry. Bidirectional aging helps reduce the CPU usage required for aging and learning MAC addresses. • Unidirectional aging mode: The system ages MAC address entries according to the source MAC address learning. Specifically, if the system does not learn a MAC address again within an aging period, the system ages the MAC address entry. For a version earlier than V800R011C00, the control board uses the unidirectional aging mode, which cannot changed to the otherwise mode.	**mac-address aging-mode**

13.3 VLAN

Virtual local area network (VLAN) is a technology used to form virtual workgroups by logically grouping the devices of a LAN.The VLAN management feature facilitates carriers' service planning.

13.3.1 Introduction

Definition

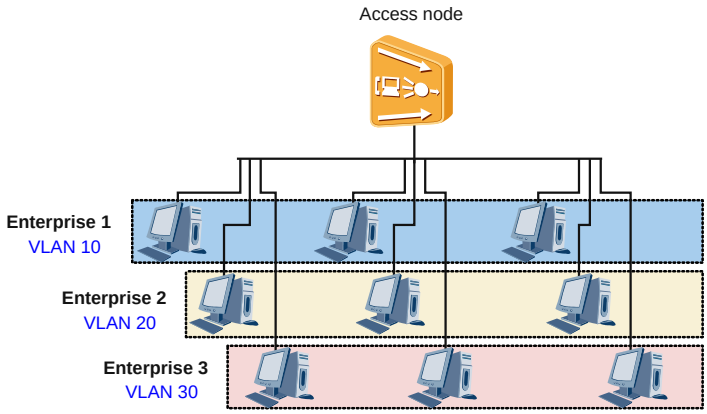
VLAN is a communications technology that divides a physical LAN into multiple logical broadcast domains (multiple VLANs). NEs in a VLAN can communicate with each other but NEs in different VLANs cannot.

Purpose

VLAN allows packets to be broadcast only within a single VLAN, preventing bandwidth waste caused by broadcast storms. Furthermore, VLAN enhances network security, because NEs in different VLANs cannot communicate with each other directly.

For example, for users of different enterprises in the same office building, it is too costly to build separate LANs for each enterprise, and insecure if the enterprises are to share the existing LAN of the building. VLAN resolves this dilemma.

Figure 13-1 VLAN application



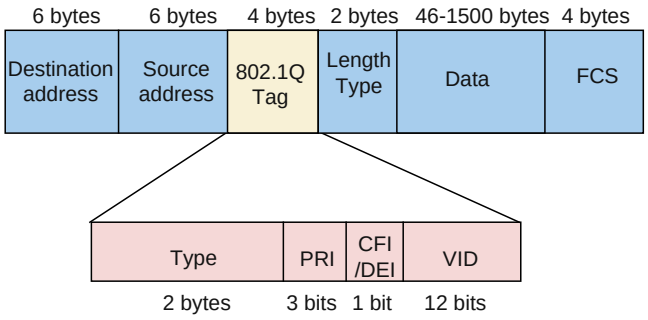
Make different enterprise users belonging to different VLANs. In this way, the enterprise users share the LAN facilities and at the same time each have their own secure networks.

13.3.2 Basic Concepts

802.1Q Frame Format for a VLAN

The format of Ethernet frames is modified in the IEEE 802.1Q standard. Specifically, a 4-byte 802.1Q tag is inserted between the source MAC address field and the protocol type field, as shown in the following figure.

Figure 13-2 802.1Q frame format for a VLAN



The 802.1Q tag consists of four fields:

- Type

  The Type field is 2 bytes long and indicates the frame type. The value 0x8100 of this field indicates that the frame carries an 802.1Q tag. Such a frame will be discarded if the receiving device does not support 802.1Q.

- PRI

  The PRI field is 3 bits long and indicates the frame priority. It ranges from 0 to 7. The larger the value, the higher the priority. In case of network congestion, the system preferentially sends data frames with a higher priority.

- CFI/DEI

  It is 1 bit long.

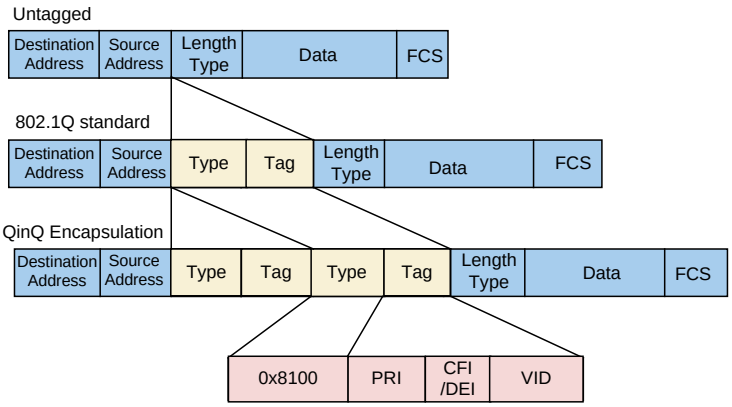
  - The CFI field indicates whether the MAC address is in canonical format. CFI 0 indicates that the MAC address is in the canonical format and CFI 1 indicates a non-canonical format. It is used to differentiate Ethernet frames, fiber distributed digital interface (FDDI) frames, and token ring frames. In an Ethernet network, CFI is 0.
  - The DEI field consists of one bit and represents the drop eligible indicator defined in the 802.1ad protocol. It is used to color the packet. For example, the value 0 means green and 1 means yellow. When the function of color-based early drop is enabled, yellow packets are dropped in case of congestion.

- VID

  VID is the shortened form of VLAN ID. It is 12 bits long and indicates the VLAN to which a frame belongs.

PRI, CFI, and VID together are called a VLAN tag, which is the basis for VLAN communication. Frames that do not contain an 802.1Q tag are called untagged frames. 802.1Q frames in QinQ encapsulation have two 802.1Q tags. The following figure shows the structures of untagged frames, standard 802.1Q frames, and 802.1Q frames in QinQ encapsulation.

Figure 13-3 Structure comparison between untagged frame, standard 802.1Q frame, and 802.1Q frame in QinQ encapsulation



VLAN Types

Table 13-5 VLAN types

VLAN Type	The same points	Difference
Standard VLAN	• Ethernet ports in different VLANs are isolated from each other. • Ethernet ports can be communicated with each other. • A VLAN contains multiple upstream ports	A standard VLAN does not contain a service port.
Smart VLAN		• A smart VLAN contains multiple service ports. • Service ports in a smart VLAN are isolated from each other.
MUX VLAN		• A MUX VLAN contains only one service port. • A service port in a MUX VLAN is isolated from a service port in another MUX VLAN. MUX VLANs and access users are in a one-to-one mapping. Therefore, a MUX VLAN uniquely identifies an access user.
Super VLAN	A super VLAN aggregates Smart VLAN or MUX VLANs. For details, see 13.3.4 VLAN Aggregation (Super VLAN).	

The Smart VLAN restricts visits between users by isolating the service flows or service ports in the same VLAN.

The MUX VLAN realizes user isolation by dividing the user service flows or service ports into different VLANs.

VLAN Attributes

Table 13-6 VLAN attributes

Attribute	Description
Common	A VLAN with the common attribute is used as a common Layer 2 VLAN. If it is used for Layer 3 forwarding, it needs a VLAN interface.
Stacking	Packets with the VLAN stacking attribute contain the inner and outer VLAN tags allocated by the MA5600T/MA5603T/MA5608T. The stacking VLAN can be used for the dual-VLAN-tag authentication on the upper-layer broadband remote access server (BRAS) and wholesaling services to Internet service providers (ISPs).
QinQ	Packets with the VLAN QinQ attribute contain the inner VLAN of the subscriber private network and the outer VLAN allocated by the MA5600T/MA5603T/MA5608T. The QinQ VLAN can be used for forming a Layer 2 virtual private network (VPN) tunnel between subscriber private networks for transparently transmitting private-network services.

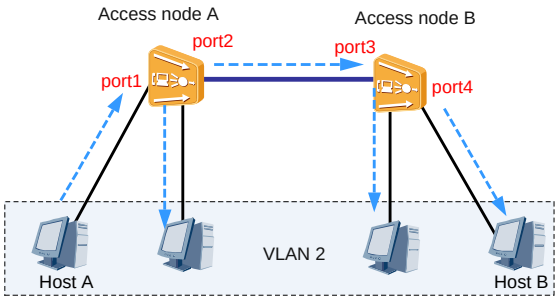
NOTE
For details about QinQ VLANs and stacking VLANs, see 13.3.5 QinQ VLAN and Stacking VLAN.

13.3.3 VLAN Communication Principle

Intra-VLAN Communication Principle

Users in a VLAN can communicate with each other. An access node implements the communication by distinguishing between the VLAN tags carried in Ethernet frames. In the following figure, there are 4 hosts in VLAN 2, and uses the example of host A and host B to illustrate communication inside a VLAN.

Figure 13-4 Intra-VLAN communication



The following describes the communication process between host A and host B inside VLAN 2.

1. Host A sends Ethernet frames to port 1 of access node A.

2. Port 1 of access node A attaches a VLAN tag (VID is filled in with VLAN 2) to the Ethernet frames.
3. Access node A sends the tagged Ethernet frames to ports (except port 1) that belong to VLAN 2.
4. Port 2 of access node A sends the tagged Ethernet frames to port 3 of access node B.
5. Access node B identifies the VLAN tag (VLAN 2) carried in the Ethernet frames and sends these frames to all ports of access node B that belong to VLAN 2.
6. Port 4 of access node B sends these Ethernet frames to host B.

Inter-VLAN Communication Principle

Hosts in different VLANs cannot communicate with each other at Layer 2. If these hosts need to communicate with each other, they need IP routes. The MA5600T/MA5603T/MA5608T, serving as a Layer 3 switch, supports communication between hosts in different VLANs.

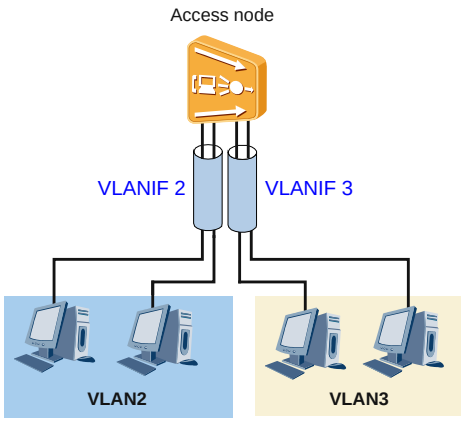
The routing table must contain the correct route entry to ensure that the first data flow is properly and correctly forwarded. In this case, a Layer 3 interface (VLAN interface) and routing protocol must be deployed on the Layer 3 switch to achieve Layer 3 route reachability. A VLAN interface is a Layer 3 logical interface and its IP address can be manually specified.

NOTE

Layer 3 switching combines routing and switching technologies to implement routing inside a switch, thereby improving the performance of the entire network. After using the routing table to send the first data flow, a Layer 3 switch generates ARP table for MAC addresses and IP addresses. When the ARP entry exists for the corresponding destination IP. The system directly find the out port by the ARP table entries and translate the destination MAC address, source MAC address, and VLAN.

In the following figure, the access node supports both VLAN 2 and VLAN 3. In this scenario, VLAN interface 2 and VLAN interface 3 can be created on the switch. After IP addressees and correct routes are configured for these two VLAN interfaces, VLAN 2 can communicate with VLAN 3.

Figure 13-5 Inter-VLAN communication



Default VLAN of an Upstream Port

The default VLAN (port default VLAN ID, or PVID) of a port is also called a native VLAN. If a default VLAN is configured on an Ethernet port, the processing rule of Ethernet port for different VLAN tag in sent packets is shown as follows.

Table 13-7 The processing rule of Ethernet port for different VLAN tag in received packets

VLAN Tag in Sent Packets on An Ethernet Port	Processing Rule of Ethernet port
Untagged	The Ethernet port tags the packets with the default VLAN and sent them.
VLAN ID is the same as the default VLAN ID	The Ethernet port removes the VLAN tag from the packets and sent them.
VLAN ID is different from the default VLAN ID	The Ethernet port directly forwards these packets with the VLAN tag unchanged and sent them.

NOTE

If the MA5600T/MA5603T/MA5608T connects to an upper-layer router, packets sent by the router may or may not carry a VLAN tag, depending on whether the router can identify a VLAN tag. The default VLAN can be configured on the upstream port of the MA5600T/MA5603T/MA5608T to adapt the router.

13.3.4 VLAN Aggregation (Super VLAN)

Context

VLAN is widely applied to switching networks because of its flexible control of broadcast domains and convenient deployment. On a Layer-3 switch, interconnection between broadcast domains is implemented using one VLAN to correspond to a single Layer-3 logical interface. However, such an implementation may lead to inefficient use of IP addresses. VLAN aggregation, also known as a super VLAN, can solve this problem.

NOTE

Here take an example of host address assignment in VLANs to show the inefficient use of IP addresses.

Example

The following figure shows a typical VLAN planning.

Figure 13-6 Typical VLAN planning

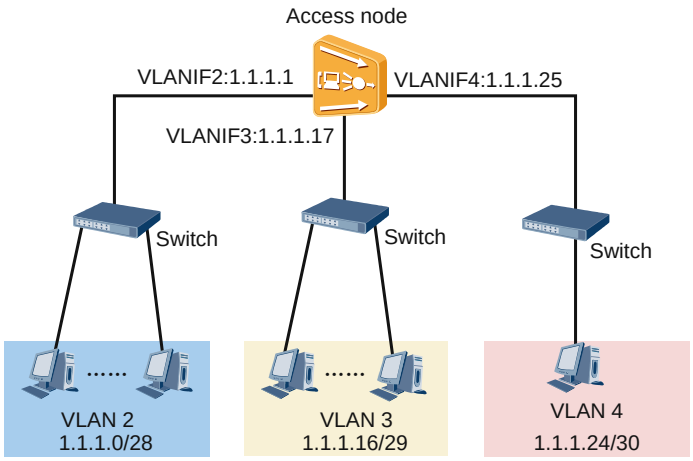


Table 13-8 Example of host address assignment in VLANs

VLAN	Subnet	Gateway Address	Number of Available Addresses	Number of Available Hosts	Actual Requirements
2	1.1.1.0/28	1.1.1.1	14	13	10
3	1.1.1.16/29	1.1.1.17	6	5	5
4	1.1.1.24/30	1.1.1.25	2	1	1

As shown in the preceding table,

- VLAN 2 requires 10 host addresses. Subnet 1.1.1.0/28 with a mask of 28 bits is assigned to VLAN 2. 1.1.1.0 is the subnet address, and 1.1.1.15 is the directed broadcast address. These two addresses cannot be used as the host address. In addition, as the default gateway address of the subnet, 1.1.1.1 cannot be used as the host address, either. The other 13 addresses ranging from 1.1.1.2 to 1.1.1.14 can be used by the hosts. In this way, although VLAN 2 needs only 10 addresses, 13 addresses need to be assigned to it according to the subnetting principle.
- VLAN 3 requires five host addresses, and subnet 1.1.1.16/29 with a mask of 29 bits needs to be assigned to VLAN 3. VLAN 4 requires only one address, and subnet 1.1.1.24/30 with a mask of 30 bits needs to be assigned to VLAN 4.

NOTE

Directed broadcast address: In a subnetted network, data packets are sent to all hosts in a subnet after a directed broadcast address is specified as the destination address of the packets. Subnet directed broadcasting is usually used to obtain data of a host in a subnet, such as the neighboring relationship.

In the above example, only 16 (10+5+1) addresses are actually required for all the VLANs. The VLAN planning, however, needs 28 (16+8+4) addresses according to the common VLAN addressing principle, even if the optimal addressing scheme is used. Therefore, nearly half of the addresses will be wasted. In addition, if VLAN 2 is accessed by only three hosts instead of ten, the remaining addresses will also be wasted because such addresses cannot be used by other VLANs.

This addressing plan is inconvenient for future network upgrades and expansion. If VLAN 4 needs an additional two hosts, but the assigned IP addresses need to remain unchanged, and the addresses after 1.1.1.24 have been assigned to other hosts, a new subnet with a mask of 29 bits and a new VLAN need to be assigned to VLAN 4's new customers. As a result, VLAN 4's customers only have three hosts, but the customers are assigned to two different subnets in separate VLANs, which makes network management difficult.

In the above example, several IP addresses are used as subnet addresses, subnet directed broadcast addresses, and default addresses of subnet gateways. Such IP addresses cannot be used as host addresses in the VLAN. This address assigning mechanism greatly reduces addressing flexibility and causes address usage waste. VLAN aggregation is developed to resolve these issues.

Principle

VLAN aggregation, divides a physical network into broadcast domains by using VLANs so that different VLANs can belong to the same subnet. VLAN aggregation consists of two basic concepts, super VLAN and sub VLAN.

- Super VLAN: Super VLANs differ from common VLANs. In super VLANs, only Layer 3 interfaces are created and physical ports are not involved. A super VLAN can be regarded as a logical Layer 3 collection of many sub VLANs.
- Sub VLAN: Sub VLANs are used to isolate broadcast domains. In sub VLANs, only physical ports are contained and Layer 3 VLAN interfaces cannot be created. The Layer 3 switching between a sub VLAN and the external network is implemented through the Layer 3 interface of the super VLAN.

A super VLAN can contain one or more sub VLANs, each sub VLAN with different broadcast domains. The sub VLAN does not occupy an independent subnet segment. In the same super VLAN, IP addresses of hosts belong to the super VLAN's subnet segment, regardless of the mapping between hosts and sub VLANs.

The same Layer 3 interface is shared by sub VLANs, allowing fewer subnet IDs, default gateway addresses of the subnet, and directed broadcast addresses of the subnet to be used. In the mean time, different broadcast domains can use the unused addresses in the same subnet segment. As such, addressing becomes flexible and efficient.

NOTE
The following still uses Table 13-8 as an example to explain the implementation principle of VLAN aggregation.

Example

Supposing that user demands are the same. In VLAN 2, ten host addresses are demanded; in VLAN 3, five host addresses are demanded; in VLAN 4, one host address is demanded.

Create VLAN 10 and configure VLAN 10 as a super VLAN. Then assign subnet address 1.1.1.0/24 with a mask of 24 bits to VLAN 10, where 1.1.1.0 is the subnet ID and 1.1.1.1 is the gateway address of the subnet, as shown in the following figure. The corresponding sub VLAN address assignment of VLAN 2, VLAN 3, and VLAN 4 is shown in Table 13-9.

Figure 13-7 VLAN aggregation schematic diagram

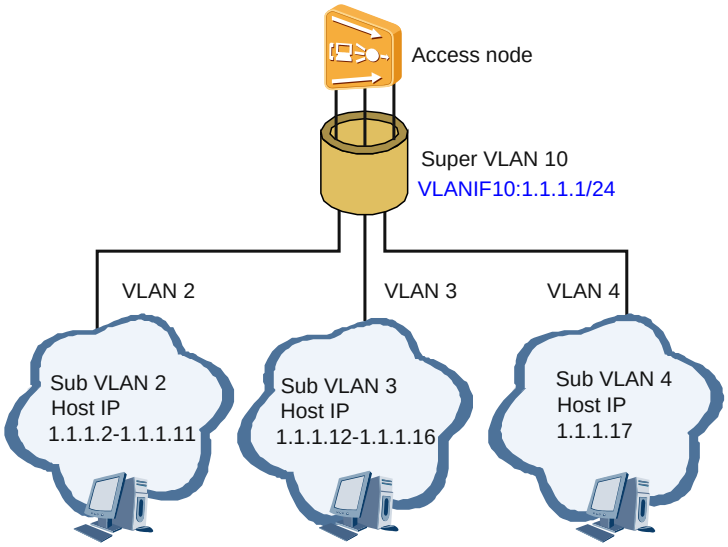


Table 13-9 Example of host address assignment in VLAN aggregation mode

VLAN	Subnet	Gateway Address	Number of Available Addresses	Available Addresses	Actual Requirements
2	1.1.1.0/24	1.1.1.1	10	1.1.1.2-1.1.1.11	10
3			5	1.1.1.12-1.1.1.16	5
4			1	1.1.1.17	1

In VLAN aggregation, sub VLANs are not divided according to the previous subnet border. Instead, their addresses are flexibly assigned among the super VLAN's subnets according to the required number of hosts. VLAN 2, VLAN 3, and VLAN 4 share a subnet (1.1.1.0/24), a default gateway address of the subnet (1.1.1.1), and a directed broadcast address of the subnet (1.1.1.255). In this manner, the subnet ID (1.1.1.16, 1.1.1.24), the default gateway of the subnet (1.1.1.17, 1.1.1.25), and the directed broadcast address of the subnet (1.1.1.5, 1.1.1.23, and 1.1.1.24) can be used as host IP addresses.

In total, 16 addresses (10+5+1) are required for the three VLANs. In practice, in this subnet, a total of 16 addresses (1.1.1.2 to 1.1.1.17) are assigned to the three VLANs. A total of 19 IP addresses are used, that is, the 16 host addresses together with the subnet ID (1.1.1.0), the default gateway of the subnet (1.1.1.1), and the directed broadcast address of the subnet (1.1.1.255). In the network segment, 236 addresses (255 - 19) are available, which can be used by any host in the sub VLAN.

Communication Between Sub VLANs

VLAN aggregation ensures that different VLANs can use the IP addresses in the same subnet segment; however, this leads to an issue with Layer 3 forwarding between sub VLANs.

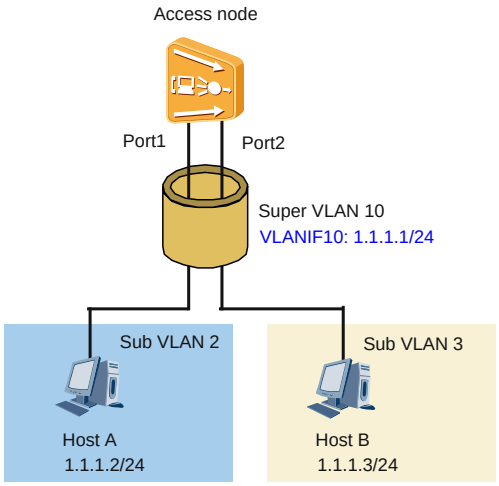
In common VLAN mode, the hosts of different VLANs can communicate with each other based on Layer 3 forwarding through their respective gateways. In VLAN aggregation mode, however, hosts in a super VLAN use IP addresses in the same network segment and share the same gateway address. Since hosts in different sub VLANs belong to the same subnet, they communicate with each other based on Layer 2 forwarding, not Layer 3 forwarding through a gateway. As a result, hosts in different sub VLANs cannot communicate with each other because the hosts are separated at Layer 2.

The Address Resolution Protocol (ARP) proxy resolves this issue. For details about the ARP proxy, see 15.3 ARP Proxy.

- **Layer 3 communication between different sub VLANs**

  As shown in the following figure, the super VLAN, namely, VLAN 10, contains two sub VLANs, VLAN 2 and VLAN 3.

  Figure 13-8 Networking diagram of Layer 3 communication between different sub VLANs based on ARP proxy



  Communication between host A in sub VLAN 2 and host B in sub VLAN 3 is implemented as follows:

  NOTE

  Suppose that host A's ARP table has no corresponding entry for host B, and the gateway between the sub VLANs is enabled with ARP proxy.

  a. After comparing the IP address (1.1.1.3) of host B with its own IP address, host A learns that both IP addresses are in the same network segment (1.1.1.0/24), and its ARP table has no entry corresponding to host B.

  b. Host A initiates ARP broadcasting to request host B's MAC address.

  c. Host B is not in the broadcast domain of sub VLAN 2, and cannot receive the ARP request.

  d. Since the gateway's ARP proxy is enabled between the sub VLANs, after receiving host A's ARP request, the gateway discovers that the IP address of host B (1.1.1.3) is the IP address of a directly-connected interface. The gateway then initiates ARP broadcasting to all other sub VLAN interfaces to request host B's MAC address.

  e. After receiving an ARP request, host B replies with an ARP response.

f. After receiving host B's ARP response, the gateway replies to host A with the gateway's MAC address. Host A will regard this MAC address as host B's MAC address.

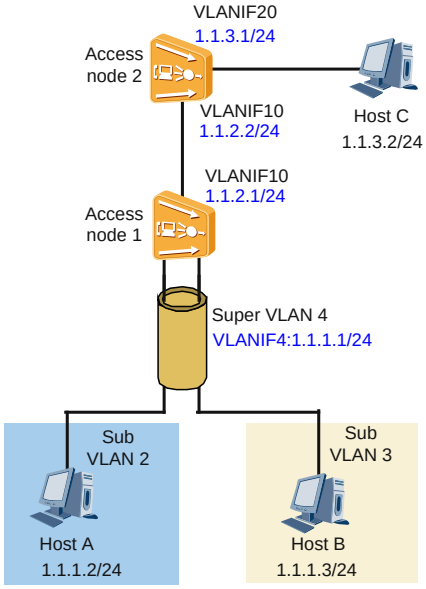
g. The ARP tables in both the gateway and host A have entries corresponding to host B.

h. To send packets to host B, host A initially sends packets to the gateway, and then the gateway performs Layer 3 forwarding to implement communication between host A and host B.

The process used by host B to send packets to host A works in the same way.

- **Layer 3 communication between a sub VLAN and an external network**

Figure 13-9 Networking diagram of Layer 3 communication between a sub VLAN and an external network



As shown in the preceding figure, access node 1 is configured with super VLAN 4, sub VLAN 2, sub VLAN 3, and common VLAN 10. Access node 2 is configured with two common VLANs, VLAN 10 and VLAN 20. Host A in sub VLAN 2 that belongs to the super VLAN 4 needs to access host C in access node 2.

NOTE

Suppose that access node 1 is configured with a route to network segment 1.1.3.0/24, and access node 2 is configured with a route to network segment 1.1.1.0/24.

a. After comparing the IP address (1.1.3.2) of host C with its IP address, host A learns that two IP addresses are not in the same network segment 1.1.1.0/24.

b. Host A initiates ARP broadcasting to its gateway, requesting the gateway's MAC address.

c. After receiving the ARP request, access node 1 identifies the correlation between the sub VLAN and the super VLAN, and replies with an ARP response to host A through sub VLAN 2. The source MAC address in the ARP response packet is the MAC address of VLANIF4 of super VLAN 4.

d. Host A learns the gateway's MAC address.

e. Host A sends the packet to the gateway, the destination MAC address being the MAC address of VLANIF4 of super VLAN 4, and the destination IP address being 1.1.3.2.

f. After receiving the packet, access node 1 performs Layer 3 forwarding and sends the packet to access node 2, with the next hop address 1.1.2.2, and the outgoing interface VLANIF10.

g. After receiving the packet, access node 2 performs Layer 3 forwarding and sends the packet to host C through the directly-connected interface VLANIF20.

h. The response packet from host C reaches access node 1 after access node 2 performs Layer 3 forwarding.

i. After receiving the packet, access node 1 performs Layer 3 forwarding and sends the packet to host A through the super VLAN.

13.3.5 QinQ VLAN and Stacking VLAN

Introduction to QinQ

As Ethernet technology is deployed in more and more carrier networks (metro Ethernet networks), the standard VLAN defined in IEEE 802.1Q alone cannot completely distinguish between and isolate a large number of users. This is because the 12-bit VLAN tag field identifies a maximum of only 4096 VLANs. The 802.1Q-in-802.1Q (QinQ) technology is developed to solve this problem.

The QinQ technology improves VLAN utilization by adding another 802.1Q tag to a packet that already carries an 802.1Q tag. With this technology, the system supports up to 4096 x 4096 VLANs. QinQ allows two VLAN headers to be inserted into a single frame. Therefore, packets transmitted in the backbone network have two 802.1Q tag headers: a public network VLAN tag and a private network VLAN tag.

For details about the format of 802.1Q frames encapsulated by QinQ, see 13.3.3 VLAN Communication Principle.

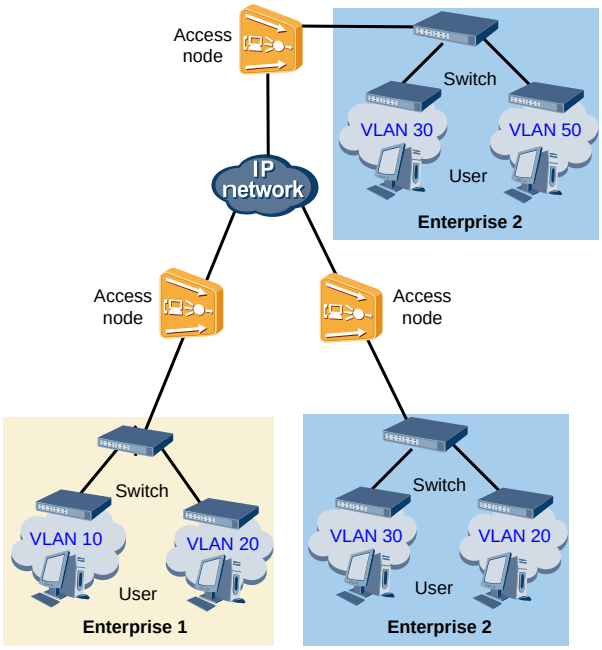
As metro Ethernet grows and a greater variety of services are deployed, there are more scenarios in which QinQ double tags can be applied.

- QinQ VLAN: Packets with two VLAN tags traverse carrier networks. The inner VLAN tag is transparently transmitted to the public network. QinQ VLAN is also a simple and practical VPN technology.
- Stacking VLAN: The inner tag indicates the user; the outer tag indicates the carrier.

QinQ VLAN

In the following network, enterprise 2 has two locations, which communicate through the carrier network (public network).

Figure 13-10 QinQ VLAN application



The access node provides the following configurations to deploy the QinQ VLAN so that enterprise 2 can communicate at two locations, but enterprise 1 and enterprise 2 cannot communicate with each other.

- The service VLAN is the same as the public network VLAN and the VLAN attribute is QinQ.
- The access node performs traffic classification for user packets based on service flows. Each user is mapped to one service flow. The inner and outer VLAN tags are configured for the service flows according to the VLAN plan in the following table.

Table 13-10 QinQ VLAN plan

Enterprise	Private Network VLAN (Inner VLAN)	Public Network VLAN (Outer VLAN)
Enterprise 1	VLANs 10 and 20	VLAN 10
Enterprise 2	VLANs 20, 30, and 50	VLAN 20

The following uses enterprise 2 users in the same private network (VLAN 30) as an example to describe packet processing:

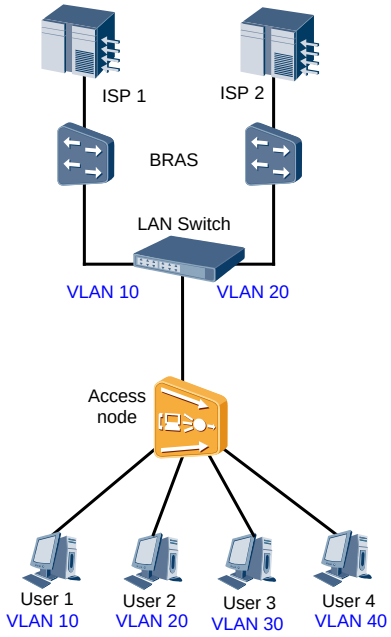
1. Untagged user packets are transmitted upstream. The packets then reach the Layer 2 switch.
2. The Layer 2 switch adds the private network VLAN tag (VLAN 30) to the packets and transmits them to the access node in the upstream direction.
3. The access node adds the public network VLAN tag (VLAN 20) to the packets and transmits the packets upstream to the upper-layer network.

4. The upper-layer network device transmits the packets to the interconnected access node according to the public network VLAN tag.
5. The interconnected access node removes the public network VLAN tag after receiving the packets, and transmits them to the switch on the same side.
6. The switch identifies and removes the private VLAN tag (VLAN 30), and forwards the untagged packets to the specified user in the private network (VLAN 30).

Stacking VLAN

The following network supports Internet service providers (ISPs) 1 and 2. User 1 and user 2 access ISP 1; user 3 and user 4 access ISP 2. The system uses the stacking VLAN to quickly provision ISP-provisioned services to specified users. Specifically, the outer VLAN tag indicates the ISP and the inner VLAN tag indicates the user.

Figure 13-11 Stacking VLAN application



The stacking VLAN is deployed according to the following configurations to connect users.

- Service VLANs with the stacking attribute are created. Two ISPs have the same outer VLAN tag.
- A service flow is created for each user.
- The inner and outer VLAN tags are configured according to the VLAN plan in the following table.

Table 13-11 Stacking VLAN plan

User	Inner VLAN Tag (Identifying Users)	Outer VLAN Tag (Identifying ISPs)
User 1	VLAN 10	VLAN 10
User 2	VLAN 20	

User	Inner VLAN Tag (Identifying Users)	Outer VLAN Tag (Identifying ISPs)
User 3	VLAN 30	VLAN 20
User 4	VLAN 40	

The following uses user 2 as an example to describe the processing of packets with the stacking VLAN from a user to an ISP:

1. User 2 sends untagged packets to the access node.
2. The access node adds two VLAN tags (inner VLAN 20 and outer VLAN 10) to the untagged packets.
3. The switch forwards the packets to ISP 1 according to outer VLAN 10.
4. ISP 1 removes outer VLAN 10 after receiving the packets and then provisions services to the user according to inner VLAN 20.

13.3.6 VLAN Translation

VLAN translation is also called VLAN mapping, which converts between user VLAN IDs and carrier VLAN IDs and allows carriers' VLAN planning to be flexible.

VLAN Translation Policies

The access node supports the following VLAN translation policies:

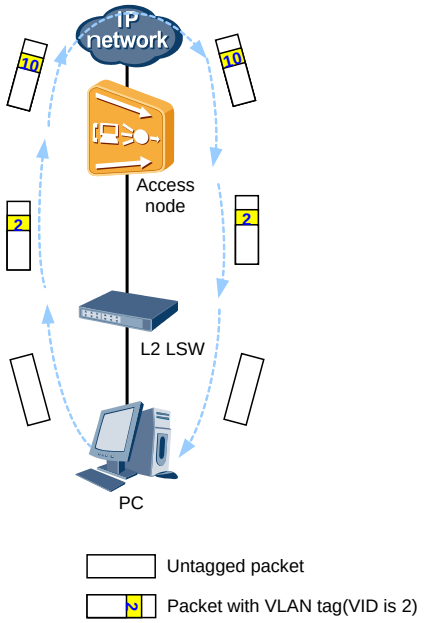
- N to 1 translation

  In the receive direction, the access node translates untagged packets, single-tagged packets, and double-tagged packets into packets with a specified VLAN tag. In the transmit direction, the access node translates this specified VLAN tag to the original VLAN tag(s) when sending packets to users.
- N to 2 translation

  In the receive direction, the access node translates untagged packets, single-tagged packets, and double-tagged packets into packets with two specified VLAN tags. In the transmit direction, the access node translates these specified VLAN tags to the original VLAN tag(s) when sending packets to users.

The following figure shows the VLAN translation process, assuming that 1 to 1 VLAN translation is used. The access node translates VLAN tag (VID 2) to VLAN tag (VID 10).

Figure 13-12 VLAN translation process



1. After receiving packets from the user side, the access node translates the VLAN tag of the packets according to the specified VLAN translation policy and then forwards the packets to the upper-layer network.
2. After receiving packets from the network side, the access node reverse translates the VLAN tag (restoring the original VLAN tag), and then sends the restored user packets to users.

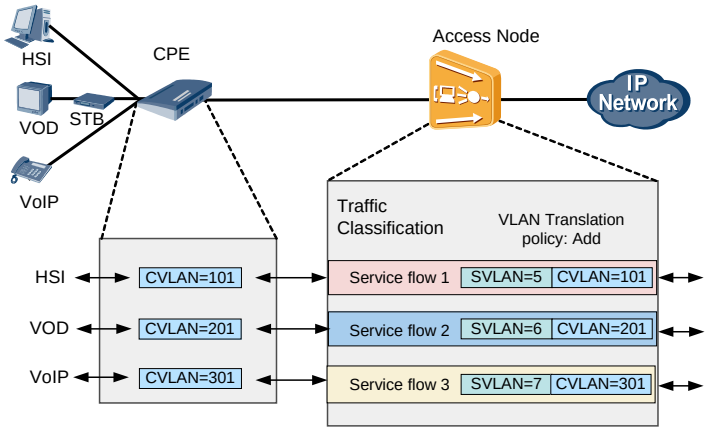
For details about VLAN translation policies and usage scenarios supported by the access node, see 13.3.8 VLAN Translation Policies Specifications.

VLAN Translation Based on Service Flow

The access node performs traffic classification for user packets based on service flows. Therefore, VLAN translation configuration and function are also implemented based on service flows.

The following figure uses VLAN translation policy "add" as an example to show the VLAN translation process on the access node.

Figure 13-13 VLAN translation based on service flow



1. The CPE attaches different C-VLAN tags to service packets.
2. The access node classifies traffic according to the C-VLAN tag and uses service flows to represent different classes.
3. The VLAN translation policy is specified to add (S-VLAN indicates the service VLAN in service flows) when service flows are configured.
4. The access node attaches an S-VLAN tag to packets matching the service flows and then forwards the packets to the upper-layer network. For example, the access node attaches VLAN tag carrying VID 5 to packets (with VLAN tag carrying VID 101) that match service flow 1.

VLANs with different attributes can be configured with different VLAN translation policies when service flows are created. For details, see 13.3.8 VLAN Translation Policies Specifications.

13.3.7 VLAN Planning Suggestion

Properly plan VLAN types and attributes to meet various service requirements.

Selecting a VLAN Type

An access device connects to users in various access modes through service flows. Select a VLAN type based on service flow requirements. A standard or super VLAN does not support the creation of service flows. Therefore, a smart or MUX VLAN is used. A MUX VLAN supports only one service flow while a smart VLAN supports multiple ones. Therefore, a smart VLAN is more popular.

A standard VLAN supports the adding of only uplink ports. Therefore, it is used for managing devices.

A super VLAN applies in Layer 3 interconnection of Layer 2 isolation. Specifically, sub-VLANs in a super VLAN can communicate with each other at Layer 3 after Address Resolution Protocol (ARP) proxy is enabled.

Selecting a VLAN Attribute

Plan the number of VLAN tags for an access device based on service requirements. For example, plan one VLAN tag for simple services and two VLAN tags for wholesale services.

One is used to identify Internet service providers (ISPs) and another is used to identify users. Then, select a VLAN attribute according to the number of VLAN tags.

- If one VLAN tag is required, select a common or QinQ VLAN.
- If two VLAN tags are required, select a stacking or QinQ VLAN.

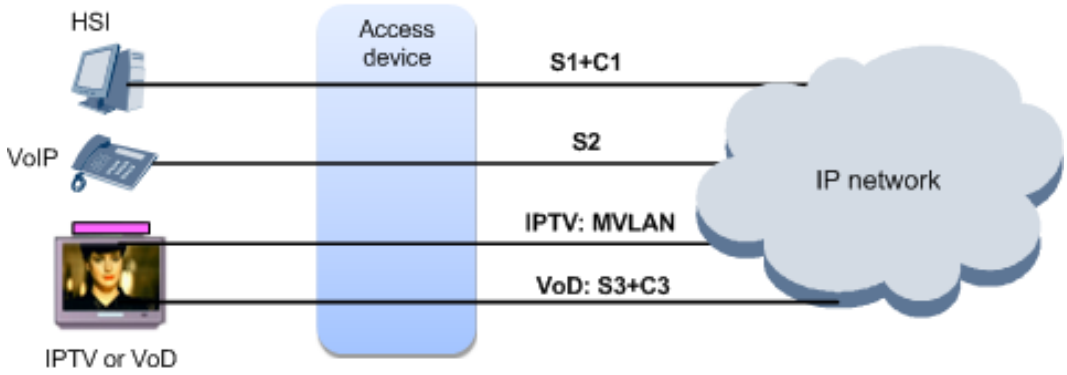
The further VLAN attribute selection rules are as follows:

- Select a QinQ VLAN for private line or transparent transmission services. The reason is that the QinQ VLAN ensures the transparent transmission of all protocol packets through hardware to the best extent.
- Select a common or stacking VLAN for enabling secure features, such as IP address anti-spoofing and MAC address anti-spoofing, Dynamic Host Configuration Protocol (DHCP) Option 82, Policy Information Transfer Protocol (PITP), and multicast functions. The reason is that a QinQ VLAN does not support the preceding functions.

VLAN Usage in the Typical Triple-play Service Scenario

In the triple-play service scenario, multiple services are provisioned for a user, such as the high-speed Internet (HSI), VoIP, IPTV, and video on demand (VoD) services.

Figure 13-14 VLAN usage in the typical triple-play service scenario



In the preceding figure:

- The HSI service uses a stacking VLAN with two VLAN tags, service VLAN (S-VLAN) tag and customer VLAN (C-VLAN) tag. S+C (S1+C1 in the preceding figure) indicates each service for each user. Enable DHCP Option 82 and PITP. Enable IP address anti-spoofing or MAC address anti-spoofing based on site requirements.
- The voice service uses a common VLAN with one VLAN tag. Enable IP address anti-spoofing and MAC address anti-spoofing.
- The VoD service uses a stacking VLAN with two VLAN tags, S-VLAN tag and C-VLAN tag (S3+C3 in the preceding figure).
- The multicast VLAN (MVLAN) of the IPTV service is a common VLAN with a VLAN tag.
- Add a VoD service flow to the MVLAN as a multicast user. The Internet Group Management Protocol (IGMP) is carried over the service flow. Therefore, the S-VLAN of the VoD service cannot be a QinQ VLAN. The reason is that a QinQ VLAN transparently transmits IGMP packets through hardware but not capture the packets. For details, see 13.3.2 Basic Concepts.

VLAN Usage in the Typical Enterprise Private Line Service Scenario

In the enterprise private line service scenario, the headquarter (HQ) and branches of an enterprise need to communicate with each other. Both the HQ and branches are in X1 VLAN.

Figure 13-15 VLAN usage in the typical enterprise private line service scenario

In the preceding figure, the VLAN tag of both access devices 1 and 2 is S1 VLAN. The X1 VLAN transparently transmits packets. DHCP or other protocols may be used in the enterprise. Therefore, use a QinQ VLAN (S1+X1 in the preceding figure) to transparently transmit the protocol packets through hardware to the best extent. This prevents global capturing and forwarding of the protocol packets through software by mistake and ensures the performance of the enterprise private line service.

13.3.8 VLAN Translation Policies Specifications

VLAN translation policies

Table 13-12 VLAN translation policies

VLAN Tag Before Translation	VLAN Tag After Translation	VLAN Translation Policy	Description	Usage Scenario
Untagged	Service VLAN (S-VLAN)	add	An S-VLAN tag is attached.	Service flows whose S-VLAN is a QinQ VLAN Transparent LAN service (TLS) flows
	Service VLAN+customer VLAN (S-VLAN+C-VLAN)	add double	Two VLAN tags, the outer S-VLAN tag and inner C-VLAN tag, are attached.	Single service flows whose S-VLAN is a stacking VLAN, and multiple service flows classified by the following rules. • untagged user-side VLAN • user-side 802.1p priority • user-side encapsulation mode Service flows whose S-VLAN is a QinQ VLAN
C-VLA	C-VLAN	transparent	The VLAN tag is transparently	Service flows whose GEM port is configured with the

VLAN Tag Before Translation	VLAN Tag After Translation	VLAN Translation Policy	Description	Usage Scenario
N			transmitted.	cascade attribute Service flows whose S-VLAN is a QinQ VLAN
	S-VLAN	translate	One VLAN tag is translated.	Multiple service flows whose S-VLAN is a common VLAN, classified by the following rules. • user-side VLAN • combination of user-side VLAN, user-side encapsulation mode • combination of user-side VLAN, user-side 802.1p priority Service flows whose S-VLAN is a QinQ VLAN
	S-VLAN+ C-VLAN	add	An S-VLAN tag is attached.	-
	S-VLAN+ C'-VLAN	translate and add	The C-VLAN tag is translated and the S-VLAN tag is attached.	Multiple service flows whose S-VLAN is a stacking VLAN, classified by the following rules. • user-side VLAN • combination of user-side VLAN, user-side encapsulation mode • combination of user-side VLAN, user-side 802.1p priority Service flows whose S-VLAN is a QinQ VLAN
C-VLAN	Untagged	remove	This is a special VLAN translation mode for connection-oriented service flows. In this mode, service flows carry one VLAN tag when arriving from the destination end, and have this VLAN tag removed when transmitted from the source end. Therefore, service flows are finally untagged.	
S-VLAN+C-VLAN	S'-VLAN +C'-VLAN	translate double	Both the S-VLAN and C-VLAN tags are translated.	Service flows whose S-VLAN is a stacking VLAN or QinQ VLAN, classified by S'-VLAN+C'-VLAN

VLAN Tag Before Translation	VLAN Tag After Translation	VLAN Translation Policy	Description	Usage Scenario
	S'-VLAN	translate and remove	The C-VLAN tag is removed and the S-VLAN tag is translated.	Service flows whose S-VLAN is a common or QinQ VLAN, classified by S-VLAN+C-VLAN
S-VLAN+C-VLAN	Untagged	remove double	This is a special VLAN translation mode for connection-oriented service flows. In this mode, service flows carry two VLAN tags (S-VLAN+C-VLAN) when arriving from the destination end, and have these VLAN tags removed when transmitted from the source end. Therefore, service flows are finally untagged.	

NOTE
Some VLAN translation policies are supported by specific boards only.

VLAN Tag Processing Policies for Switch-oriented service Flows with Different VLAN Attributes

The following tables describe the default and configurable VLAN translation policies and the corresponding VLAN tag processing policies for switch-oriented service flows with different VLAN attributes.

NOTE
- S-Tag: S-VLAN Tag, indicates the service VLAN tag.
- C-Tag: C-VLAN Tag, indicates the user VLAN tag.
- S'-Tag: S'-VLAN Tag, indicates another service VLAN tag.
- C'-Tag: C'-VLAN Tag, indicates another user VLAN tag.
- Any: Any VLAN Tag
- untagged: No VLAN tag
- IPoE: Include IPv4oE and IPv6oE

Table 13-13 VLAN translation policies for switch-oriented service flows whose S-VLAN is a QinQ VLAN

Traffic Classification Type	Traffic Classification Parameter	Default VLAN Translation Policy	Configurable VLAN Translation Policy	Tag Processing Policy for Upstream Packets
Single service	NA	add	add	Any → S-Tag+Any
			add double	Any → S-Tag+C-Tag+Any
User-side	PPPoE	add	add	Any → S-Tag+Any

Traffic Classification Type	Traffic Classification Parameter	Default VLAN Translation Policy	Configurable VLAN Translation Policy	Tag Processing Policy for Upstream Packets
encapsulation mode			add double	Any → S-Tag+C-Tag+Any
	IPoE	add	add	Any → S-Tag+Any
			add double	Any → S-Tag+C-Tag+Any
User-side VLAN	C-VLAN: 1-4095	add	add	C-Tag → S-Tag+C-Tag
			add double	C-Tag → S-Tag+C'-Tag+C-Tag
			transparent	C-tag → S-tag (C-tag = S-tag)
			translate	C-Tag → S-Tag
			translate and add	C-Tag → S-Tag+C'-Tag
	priority-tagged	add	add	pri-tag → S-Tag+pri-tag
			add double	pri-tag → S-Tag+C-Tag+pri-tag
			translate	pri-tag → S-Tag
			translate and add	pri-tag → S-Tag+C'-Tag
	untagged	add	add	untagged → S-Tag
			add double	untagged → S-Tag+C-Tag
	other-all (any other)	add	add	Any → S-Tag+Any
			add double	Any → S-Tag+C-Tag+Any
	VLAN range	add	add	C-VLAN range → S-Tag+C-VLAN range
User-side VLAN+user-side 802.1p priority	VLAN: 1-4095 P-bits: 0-7	add	add	C-Tag → S-Tag+C-Tag
			add double	C-Tag → S-Tag+C'-Tag+C-Tag
			transparent	C-tag → S-tag (C-tag = S-tag)
			translate	C-Tag → S-Tag
			translate and add	C-Tag → S-Tag+C'-Tag
	VLAN: priority-tagged	add	add	pri-tag → S-Tag+pri-tag
			add double	pri-tag → S-Tag+C-Tag+pri-tag

Traffic Classification Type	Traffic Classification Parameter	Default VLAN Translation Policy	Configurable VLAN Translation Policy	Tag Processing Policy for Upstream Packets
	P-bits: 0-7		translate	pri-tag → S-Tag
			translate and add	pri-tag → S-Tag+C'-Tag
User-side VLAN+user-side encapsulation mode	VLAN: 1-4095 User-side encapsulation mode: PPPoE/IPoE	add	add	C-Tag → S-Tag+C-Tag
			add double	C-Tag → S-Tag+C'-Tag+C-Tag
			transparent	C-tag → S-tag (C-tag = S-tag)
			translate	C-Tag → S-Tag
			translate and add	C-Tag → S-Tag+C'-Tag
	VLAN: priority-tagged User-side encapsulation mode: PPPoE/IPoE	add	add	pri-tag → S-Tag+pri-tag
			add double	pri-tag → S-Tag+C-Tag+pri-tag
			translate	pri-tag → S-Tag
			translate and add	pri-tag → S-Tag+C'-Tag
	VLAN: untagged User-side encapsulation mode: PPPoE/IPoE	add	add	untagged → S-Tag
			add double	untagged → S-Tag+C-Tag
Two VLAN tags (S-VLAN+C-VLAN)	Outer VLAN: 1-4095 Inner VLAN: 1-4095	translate and remove	translate double	S-Tag+C-Tag → S'-Tag+C'-Tag
			translate and remove	S-Tag+C-Tag → S'-Tag

Table 13-14 VLAN translation policies for switch-oriented service flows whose S-VLAN is a common VLAN

NOTE
Only some boards support bold VLAN translation policy.

Traffic Classification Type	Traffic Classification Parameter	Default VLAN Translation Policy	Configurable VLAN Translation Policy	Tag Processing Policy for Upstream Packets
Single service	NA	add	add	untagged → S-Tag pri-tag → S-Tag
			add double	**untagged → S-Tag+C-Tag** **pri-tag → S-Tag+C-Tag**
User-side encapsulation mode	PPPoE	add	add	untagged → S-Tag pri-tag → S-Tag
			add double	**untagged → S-Tag+C-Tag** **pri-tag → S-Tag+C-Tag**
	IPoE	add	add	untagged → S-Tag pri-tag → S-Tag
			add double	**untagged → S-Tag+C-Tag** **pri-tag → S-Tag+C-Tag**
User-side VLAN	C-VLAN: 1-4095	translate	**add**	**C-Tag → S-Tag+C-Tag**
			add double	**C-Tag → S-Tag+C'-Tag+C-Tag**
			transparent	C-Tag → S-Tag (C-tag = S-tag, GEM port cascading)
			translate	C-Tag → S-Tag
			translate and add	**C-Tag → S-Tag+C'-Tag**
	priority-tagged	translate	**add**	**pri-tag → S-Tag**
			add double	**pri-tag → S-Tag+C-Tag**
			translate	pri-tag → S-Tag
			translate and add	**pri-tag → S-Tag+C'-Tag**
	untagged	add	add	untagged → S-Tag
			add double	**untagged → S-Tag+C-Tag**
	other-all (any other)	add	**add**	**Any → S-Tag+Any**
			add double	**Any → S-Tag+C-Tag+Any**
	VLAN range	add	**add**	**C-Tag range → S-Tag+C-Tag range**
User-side	VLAN: 1-4095	translate	**add**	**C-Tag → S-Tag+C-Tag**

Traffic Classification Type	Traffic Classification Parameter	Default VLAN Translation Policy	Configurable VLAN Translation Policy	Tag Processing Policy for Upstream Packets
VLAN+user-side 802.1p priority	P-bits: 0-7		**add double**	**C-Tag → S-Tag+C'-Tag+C-Tag**
			transparent	C-Tag → S-Tag (C-tag = S-tag, GEM port cascading)
			translate	C-Tag → S-Tag
			translate and add	**C-Tag → S-Tag+C'-Tag**
	VLAN: priority-tagged P-bits: 0-7	translate	**add**	**pri-tag → S-Tag**
			add double	**pri-tag → S-Tag+C-Tag**
			translate	pri-tag → S-Tag
			translate and add	**pri-tag → S-Tag+C'-Tag**
User-side VLAN+user-side encapsulation mode	VLAN: 1-4095 User-side encapsulation mode: PPPoE/IPoE	translate	**add**	**C-Tag → S-Tag+C-Tag**
			add double	**C-Tag → S-Tag+C'-Tag+C-Tag**
			transparent	C-Tag → S-Tag (C-tag = S-tag, GEM port cascading)
			translate	C-Tag → S-Tag
			translate and add	**C-Tag → S-Tag+C'-Tag**
	VLAN: priority-tagged User-side encapsulation mode: PPPoE/IPoE	translate	**add**	**pri-tag → S-Tag**
			add double	**pri-tag → S-Tag+C-Tag**
			translate	pri-tag → S-Tag
			translate and add	**pri-tag → S-Tag+C'-Tag**
	VLAN: untagged User-side encapsulation mode: PPPoE/IPoE	add	add	untagged → S-Tag
			add double	**untagged → S-Tag+C-Tag**
Two VLAN tags (S-VLAN+	outer VLAN: 1-4095 inner VLAN: 1-4095	translate and remove	**translate double**	**S-Tag+C-Tag → S'-Tag+C'-Tag**
			translate and	S-Tag+C-Tag → S'-Tag

Traffic Classification Type	Traffic Classification Parameter	Default VLAN Translation Policy	Configurable VLAN Translation Policy	Tag Processing Policy for Upstream Packets
C-VLAN)			remove	

Table 13-15 VLAN translation policies for switch-oriented service flows whose S-VLAN is a stacking VLAN

Traffic Classification Type	Traffic Classification Parameter	Default VLAN Translation Policy	Configurable VLAN Translation Policy	Tag Processing Policy for Upstream Packets
Single service	NA	add double	add double	untagged → S-Tag+C-Tag pri-tag → S-Tag+C-Tag C-Tag → S-Tag+C'-Tag+C-Tag
User-side encapsulation mode	PPPoE	add double	add double	untagged → S-Tag+C-Tag pri-tag → S-Tag+C-Tag C-Tag → S-Tag+C'-Tag+C-Tag
	IPoE	add double	add double	untagged → S-Tag+C-Tag pri-tag → S-Tag+C-Tag C-Tag → S-Tag+C'-Tag+C-Tag
User-side VLAN	C-VLAN: 1-4095	transparent	transparent	C-tag → S-tag (C-tag = S-tag, GEM port cascading)
		translate and add	translate and add	C-Tag → S-Tag+C'-Tag
	priority-tagged	translate and add	translate and add	pri-tag → S-Tag+C'-Tag
	untagged	add double	add double	untagged → S-Tag+C-Tag
	other-all (any other)	add	add	Any → S-Tag + Any
			add double	Any → S-Tag +C-Tag+ Any
	VLAN range	add	add	C-VLAN range →

Traffic Classification Type	Traffic Classification Parameter	Default VLAN Translation Policy	Configurable VLAN Translation Policy	Tag Processing Policy for Upstream Packets
				S-Tag+C-VLAN range
User-side VLAN + 802.1p priority	VLAN: 1-4095 P-bits: 0-7	transparent	transparent	C-tag → S-tag (C-tag = S-tag, GEM port cascading)
		translate and add	translate and add	C-Tag → S-Tag+C'-Tag
	VLAN: priority-tagged P-bits: 0-7	translate and add	translate and add	pri-tag → S-Tag+C'-Tag
User-side VLAN+user-side encapsulation mode	VLAN: 1-4095 User-side encapsulation mode: PPPoE/IPoE	transparent	transparent	C-tag → S-tag (C-tag = S-tag, GEM port cascading)
		translate and add	translate and add	C-Tag → S-Tag+C'-Tag
	VLAN: priority-tagged User-side encapsulation mode: PPPoE/IPoE	translate and add	translate and add	pri-tag → S-Tag+C'-Tag
	VLAN: untagged User-side encapsulation mode: PPPoE/IPoE	add double	add double	untagged → S-Tag+C-Tag
Two VLAN tags (S-VLAN +C-VLAN)	Outer VLAN: 1-4095 Inner VLAN: 1-4095	translate double	translate double	S-Tag+C-Tag → S'-Tag+C'-Tag

13.3.9 Configuring a VLAN

A virtual local area network (VLAN) is used to separate broadcast domains. VLANs have enhanced security and support expansion and flexible networking. Configuring VLAN is a prerequisite for configuring a service. Hence, before configuring a service, make sure that the VLAN configuration based on planning is complete.

Application Context

VLAN application is specific to user types. For details on the VLAN application, see Table 13-16.

Table 13-16 VLAN application and planning

Application Scenario	VLAN Planning	Remarks
The device communicates with a network management system (NMS) or the devices are cascaded.	VLAN type: standard VLAN VLAN attribute: common VLAN forwarding mode: VLAN+MAC	This plan is only applicable to Ethernet ports. A standard VLAN only contains multiple uplink ports. The Ethernet ports in one VLAN can communicate with each other and the Ethernet ports in different VLANs are separated.
N:1 access	VLAN type: smart VLAN attribute: common VLAN forwarding mode: by VLAN+MAC	This plan is used for converging and transmitting the services of multiple users to one VLAN. This reduces the number of VLANs to be used. Smart VLANs can be applied in residential communities to provide xDSL or xPON service access.
1:1 access • Multi-ISP wholesale service (ISP is the abbreviation of Internet service provider.) • VLAN ID expansion service	VLAN type: smart Attribute: stacking VLAN forwarding mode: by S+C	A stacking VLAN packet contains two VLAN tags (inner and outer VLAN) assigned by the MA5600T/MA5603T/MA5608T. The outer VLAN identifies a service (service VLAN) and the inner VLAN identifies a user (customer VLAN). The service used by a user is identified by a unique VLAN ID (S+C VLAN). The upper layer broadband remote access server (BRAS) authenticates a service by the two VLAN tags contained in the service packets. In this manner, the BRAS can receive packets from more users. On a Layer-2 upper layer network of the MA5600T/MA5603T/MA5608T, packets can also be forwarded by the outer VLAN+MAC ID. This function enables ISPs to provide the wholesale services.

Application Scenario	VLAN Planning	Remarks
Enterprise private line service access	VLAN type: smart VLAN attribute: QinQ VLAN forwarding mode: by VLAN+MAC or S+C.	A QinQ VLAN contains an inner VLAN tag from the user private network, and an outer VLAN tag assigned by the MA5600T/MA5603T/MA5608T. The outer VLAN forms a Layer 2 virtual private network (VPN) on the user private network for transparently transmitting user services on the private network.

Creating a VLAN

Creating VLAN is a prerequisite for configuring a VLAN or service. Hence, before configuring a VLAN, make sure that the VLAN creating based on planning is complete.

Prerequisites

The ID of the planned VLAN is not occupied.

Application Context

VLAN application is specific to user types. For details on the VLAN application, see Table 13-17.

Table 13-17 VLAN planning

User Type	Application Scenario	VLAN Planning
• Household user • Commercial user of the Internet access service	N:1 scenario, that is, the scenario of upstream transmission through a single VLAN, where the services of multiple users are converged to the same VLAN.	VLAN type: smart
	1:1 scenario, that is, the scenario of upstream transmission through double VLANs, where the outer VLAN tag identifies a service and the inner VLAN tag identifies a user. The service of each user is indicated by a unique S+C.	
Commercial	Applicable only to the	

User Type	Application Scenario	VLAN Planning
user of the transparent transmission service	transparent transmission service of a commercial user.	

Default Configuration

Table 13-18 lists the default parameter settings of VLAN.

Table 13-18 Default parameter settings of VLAN

Parameter	Default Setting	Remarks
Default VLAN of the system	VLAN ID: 1 Type: smart VLAN	You can run the **defaultvlan modify** command to modify the VLAN type but cannot delete the VLAN.
Reserved VLAN of the system	VLAN ID range: 4079-4093	You can run the **vlan reserve** command to modify the VLAN reserved by the system.

Prerequisite

- The VLAN to be added should not exist in the system.
- Service VLAN cannot be reserve VLAN.

Procedure

Create a VLAN.

Run the **vlan** to create a VLAN. VLANs of different types are applicable to different scenarios.

Table 13-19 VLAN types and application scenarios

VLAN Type	Configuration Command	VLAN Description	Application Scenario
Standard VLAN	To add a standard VLAN, run the **vlan *vlanid* standard** command.	Standard VLAN. Ethernet ports in a standard VLAN are interconnected with each other but Ethernet ports in different standard VLANs are isolated from each other.	Only available to Ethernet ports and specifically to network management and subtending.
Smart VLAN	To add a smart VLAN, run the **vlan *vlanid***	One VLAN may contain multiple xDSL service ports	Smart VLANs can be applied in residential communities to provide xDSL or xPON

VLAN Type	Configuration Command	VLAN Description	Application Scenario
	smart command.	or xPON service ports. The traffic streams of these ports, however, are isolated from each other. In addition, the traffic streams of different VLANs are also isolated. One smart VLAN provides access for multiple users and therefore saves VLAN resources.	service access.
MUX VLAN	To add a MUX VLAN, run the **vlan *vlanid* mux** command.	One MUX VLAN contains only one xDSL service port or xPON service port. The traffic streams in different VLANs are isolated from each other. One-to-one mapping can be set up between a MUX VLAN and an access user. Hence, a MUX VLAN can identify an access user.	MUX VLANs are applicable to xDSL or xPON service access. For example, MUX VLANs can be used to distinguish users.
Super VLAN	To add a super VLAN, run the **vlan *vlanid* super** command.	The super VLAN is based on Layer 3. One super VLAN contains multiple sub-VLANs. Through an ARP proxy, the sub-VLANs in a super VLAN can be interconnected at Layer 3.	Super VLANs save IP addresses and improve the utilization of IP addresses. For a super VLAN, sub-VLANs must be configured. You can run the **supervlan** command to add a sub-VLAN to a specified super VLAN. A sub-VLAN must be a smart VLAN or MUX VLAN.

NOTE

- To add VLANs with consecutive IDs in batches, run the **vlan *vlanid* to *end-vlanid*** command.
- To add VLANs with inconsecutive IDs in batches, run the **vlan *vlan-list*** command.

----End

Example

Create VLAN 50 for extension of the VLAN ID. And the type of VLAN is smart.

```
huawei(config)#vlan 50 smart
```

Create VLAN 55-60 for extension of the VLAN ID. And the type of VLAN is smart.

```
huawei(config)#vlan 55 to 60 smart
```

Create VLAN 65, 73 and 52 for extension of the VLAN ID. And the type of VLAN is smart.

```
huawei(config)#vlan 65,73,52 smart
```

Configuring the VLAN attribute

Configuring the VLAN attribute is a prerequisite for configuring a VLAN. Hence, before configuring a service, make sure that the VLAN attribute configuration based on planning is complete.

Application Context

VLAN application is specific to user types. For details on the VLAN application, see Table 13-20.

Table 13-20 VLAN attribute planning

User Type	Application Scenario	VLAN Planning
• Household user • Commercial user of the Internet access service	N:1 scenario, that is, the scenario of upstream transmission through a single VLAN, where the services of multiple users are converged to the same VLAN.	VLAN attribute: common
	1:1 scenario, that is, the scenario of upstream transmission through double VLANs, where the outer VLAN tag identifies a service and the inner VLAN tag identifies a user. The service of each user is indicated by a unique S+C.	Attribute: stacking
Commercial user of the transparent transmission service	Applicable only to the transparent transmission service of a commercial user.	VLAN attribute: QinQ

Default Configuration

Table 13-21 lists the default parameter settings of VLAN.

Table 13-21 Default attribute settings of VLAN

Parameter	Default Setting
Default attribute of a new VLAN	Common

Prerequisite

- The VLAN to be configured should have been created.
- The VLAN attribute must be planned properly according to the application scenarios.

Procedure

Configure the VLAN attribute.

The default attribute for a new VLAN is "common". You can run the **vlan attrib** command to configure the attribute of the VLAN.

Configure the attribute according to VLAN planning.

Table 13-22 VLAN attributes and application scenarios

VLAN Attribute	Configuration Command	VLAN Type	VLAN Description	Application Scenario
Common	The default attribute for a new VLAN is "common".	The VLAN with this attribute can be a standard VLAN, smart VLAN, MUX VLAN, or super VLAN.	A VLAN with the common attribute can function as a common Layer 2 VLAN or function for creating a Layer 3 interface.	Applicable to the N:1 access scenario.
QinQ VLAN	To configure QinQ as the attribute of a VLAN, run the **vlan attrib *vlanid* q-in-q** command.	The VLAN with this attribute can be a standard VLAN, smart VLAN or MUX VLAN. The attribute of a sub VLAN, the VLAN with a Layer 3 interface, and the default VLAN of the system cannot be	The packets from a QinQ VLAN contain two VLAN tags, that is, inner VLAN tag from the private network and outer VLAN tag from the MA5600T/MA5603T/MA5608T. Through the outer VLAN, a	Applicable to the enterprise private line scenario. A QinQ VLAN does not support the following services: DHCP Option 82, DHCP Layer 2 relay, DHCP Layer 3 relay, PITP P mode, IP address

VLAN Attribute	Configuration Command	VLAN Type	VLAN Description	Application Scenario
		set to QinQ VLAN.	Layer 2 VPN tunnel can be set up to transparently transmit the services between private networks.	anti-spoofing, MAC address anti-spoofing, multicast IPTV, and ARP proxy.
VLAN Stacking	To configure stacking as the attribute of a VLAN, run the **vlan attrib *vlanid* stacking** command.	The VLAN with this attribute can only be a smart VLAN or MUX VLAN. The attribute of a sub VLAN, the VLAN with a Layer 3 interface, and the default VLAN of the system cannot be set to VLAN stacking.	The packets from a stacking VLAN contain two VLAN tags, that is, inner VLAN tag and outer VLAN tag from the MA5600T/MA5603T/MA5608T. The upper-layer BRAS authenticates the access users according to the two VLAN tags. In this manner, the number of access users is increased. On the upper-layer network in the Layer 2 working mode, a packet can be forwarded directly by the outer VLAN tag and MAC address mode to provide the wholesale service for ISPs.	Applicable to the 1:1 access scenario for the wholesale service or extension of VLAN IDs. In the case of a stacking VLAN, to configure the inner tag of the service port, run the **stacking label** command.

NOTE

- To configure attributes for the VLANs with consecutive IDs in batches, run the **vlan attrib** *vlanid* to *end-vlanid* command.
- To configure attributes for the VLANs with inconsecutive IDs in batches, run the **vlan attrib** *vlan-list* command.

----End

Example

To configure the attribute of VLAN 50 to **stacking** for extending VLAN IDs, do as follows:

```
huawei(config)#vlan attrib 50 stacking
```

To configure the attributes of VLANs 55-60 (used for enterprise users) to **QinQ**, do as follows:

```
huawei(config)#vlan attrib 55 to 60 q-in-q
```

To configure the attributes of service VLANs 65, 73, and 52 to **stacking**, do as follows:

```
huawei(config)#vlan attrib 65,73,52 stacking
```

Configuring an Upstream Port

The uplink port on an access device connects to the upper layer device to forward access device data to the upstream network and forward upper layer device data to users.

Prerequisites

The planned virtual local area network (VLAN) is already configured.

Procedure

Configure an upstream port for the VLAN.

Run **port vlan** command to add the upstream port to the VLAN.

Step 1 Configure the attribute of the upstream port.

If the default attribute of the upstream port does not meet the requirement for interconnection of the upstream port with the upper-layer device, you need to configure the attribute. For configuration details, see Configuring the Attributes of an Upstream Ethernet Port.

Step 2 (Optional) Configure redundancy backup for the uplink.

To ensure reliability of the uplink, two upstream ports must be available. That is, redundancy backup of the upstream ports needs to be configured. For details, see 19.3.5 Configuring Ethernet Link Aggregation.

----End

Example

Assume that the 0/19/0 and 0/19/1 upstream ports are to be added to VLAN 50. The 0/19/0 and 0/19/1 need to be configured into an aggregation group for double upstream accesses. For the two upstream ports, the working mode is full-duplex (full) and the port rate is 100 Mbit/s. To configure such upstream ports, do as follows:

```
huawei(config)#port vlan 50 0/19 0
huawei(config)#port vlan 50 0/19 1
huawei(config)#interface giu 0/19
huawei(config-if-giu-0/19)#duplex 0 full
huawei(config-if-giu-0/19)#duplex 1 full
huawei(config-if-giu-0/19)#speed 0 100
huawei(config-if-giu-0/19)#speed 1 100
```

```
huawei(config-if-giu-0/19)#quit
huawei(config)#link-aggregation 0/19 0 0/19 1 egress-ingress workmode lacp-static
```

Configuring the Upstream VLAN Interface

This topic describes how to specify the upstream VLAN interface for the media stream and the signaling flow, and how to configure the IP addresses of the Layer 3 interface. These IP addresses are the sources of the media and signaling IP address pools.

Context

Multiple IP addresses can be configured for the same VLAN Layer 3 interface. Only one IP address functions as the primary address, and other IP addresses function as the secondary addresses.

Procedure

Run the **vlan** command to add an upstream VLAN for the media stream and the signaling flow.

Step 1 Run the **port vlan** command to add the upstream ports to the VLAN.

Step 2 Configure the IP addresses of the VLAN Layer 3 interface.

1. Run the **interface vlanif** command to enter the Layer 3 interface of the upstream VLAN for the media stream and the signaling flow.
2. Run the **ip address** command to configure the IP addresses of the Layer 3 interface.

Step 3 Run the **display interface vlanif** command to check whether the IP addresses of the Layer 3 interface are the same as those in the data plan.

----End

Example

Assume that the media stream and the signaling stream are transmitted upstream through upstream port 0/19/0. To create media and signaling upstream VLAN 3 and configure IP addresses 10.13.4.116/16 and 10.13.4.117/16 of the Layer 3 interfaces for the media IP address pool and the signaling IP address pool respectively, do as follows:

```
huawei(config)#vlan 3 standard
huawei(config)#port vlan 3 0/19 0
huawei(config)#interface vlanif 3
huawei(config-if-vlanif3)#ip address 10.13.4.116 16
huawei(config-if-vlanif3)#ip address 10.13.4.117 16 sub
huawei(config-if-vlanif3)#quit
huawei(config)#display interface vlanif 3
vlanif3 current state : UP
Line protocol current state : UP
Description : HUAWEI, SmartAX Series, vlanif3 Interface
The Maximum Transmit Unit is 1500 bytes
Internet Address is 10.13.4.116/16
Internet Address is 10.13.4.117/16 Secondary
IP Sending Frames' Format is PKTFMT_ETHNT_2, Hardware address is 00e0-fc32-1118
```

Configuring a VLAN Service Profile

A VLAN service profile is a collection of service-related parameters for VLAN attributes. After a VLAN is bound to a VLAN service profile, the VLAN has all the VLAN attributes defined in the VLAN service profile. Binding a VLAN service profile is an efficient way of configuring a VLAN.

Application Context

VLAN, as a basic and also important concept of access equipment, involves discrete configurations of many parameters. These parameters include forwarding mode, security feature, protocol enabling/disabling, transparent transmission of protocol packets, and packet forwarding policy. Service parameters are related to specific VLANs and in actual usage there are a lot of VLANs, causing complex configuration. Against this backdrop, the VLAN service profile is introduced to achieve simplified and highly-efficient configuration. A VLAN service profile is abstracted from specific VLANs and supports centralized configuration of VLAN-related service parameters. Different VLANs of the same attribute can flexibly be bound to (or unbound from) a VLAN service profile to possess (or release) the attributes defined in the VLAN service profile.

Prerequisite

The VLAN to which the VLAN service profile is bound must be created.

Configuration Process

1. Create a VLAN service profile.
2. Configure the following service parameters according to service requirements:
   - Forwarding mode
   - Forwarding policy
   - Protocol switch
   - Transparent transmission function
   - Security function
3. Commit to save the current parameters.
4. Bind the VLAN service profile to the VLAN.

Procedure

Create a VLAN service profile.

Run the **vlan service-profile** command to create a VLAN service profile or enter the configuration mode of the VLAN service profile. When the profile does not exist, running this command means to create a VLAN service profile and enter the configuration mode of the service profile. If the profile exists, running this command means to directly enter the configuration mode of this service profile.

Step 1 Configure the VLAN forwarding mode.

Forwarding mode refers to the Layer 2 packet forwarding mechanism, including VLAN+MAC forwarding (default) and SVLAN+CVLAN (or S+C) forwarding. In VLAN+MAC forwarding, the system needs to dynamically learn the mapping relationship between VLAN, source MAC address, and port. In S+C forwarding, the system does not need to dynamically learn MAC addresses but determines the forwarding entry according to

SVLAN and CVLAN. Because S+C forwarding does not depend on MAC address learning, it has the following advantages:

1. Saving MAC addresses
2. Preventing occurrence of unknown unicast packets caused by aging of dynamic MAC addresses Broadcasting unknown unicast packets threatens the security of the device
3. Ensuring security by solving problems such as MAC spoofing and attack

- Run the **forwarding** command to configure the VLAN forwarding policy.
- Run the **user-bridging** command to configure the bridging function of the VLAN service profile. After the bridging function is enabled, two users in the same VLAN can directly communicate with each other at Layer 2.

NOTE
The bridging function is visible only to the SCUN/SCUK control board. It conflicts with S+C forwarding.

- Run the **mac-address learning** command to configure MAC address learning on the control board.

Step 2 Configure the VLAN forwarding policy.

Forwarding policy refers to the discard policy of packets such as downstream broadcast, downstream unknown unicast, and unknown multicast packets.

- Run the **packet-policy** command to configure the forwarding policy for the downstream broadcast packets, downstream unknown unicast packets, and unknown multicast packets in the VLAN. Two policies namely forward and discard are supported.
- Run the **igmp mismatch** command to configure the mismatch IGMP policy of the VLAN, supports the transparent and discard policies.

Step 3 Configure the VLAN protocol switch.

Protocol switch refers to whether to enable certain types of protocols or certain functions of a protocol. VMAC aging and PPPoE MAC conflict with S+C forwarding.

- Run the **dhcp mode** command to configure the DHCPv4 forwarding mode, that is, to switch between the DHCP Layer 2 forwarding mode and the DHCP Layer 3 forwarding mode.
- Run the **dhcpv6 mode** command to configure the DHCPv6 forwarding mode, that is, to switch between the DHCP Layer 2 forwarding mode and the DHCP Layer 3 forwarding mode.
- Run the **pppoe mac-mode** command to configure the MAC address allocation mode for PPPoE users. Two modes namely, single-mac and multi-mac are supported.
- Run the **pppoa mac-mode** command to configure the MAC address allocation mode for PPPoA users. Two modes namely, single-mac and multi-mac are supported.
- Run the **vmac aging-mode** command to configure the VMAC aging mode, which can be common aging or DHCP-based aging.
- Run the **pitp** command to configure the PITP function to implement authentication of bound user account and access port.
- Run the **dhcp option82** command to configure the DHCPv4 option 82 feature.
- Run the **dhcpv6 option82** command to enable or disable the DHCPv6 option feature.
- Run the **dhcp proxy** command to configure the DHCP proxy function. After the DHCP proxy function is enabled, the server ID proxy function and lease time proxy function will be enabled.

Step 4 Configure the VLAN transparent transmission function.

In transparent transmission, the system does not process specified types of protocol packets but transparently transmits them.

- Run the **bpdu tunnel** command to configure the BPDU transparent transmission switch. After transparent transmission is enabled, the Layer 2 BPDUs of the private network can be transmitted transparently over the public network.
- Run the **vtp-cdp tunnel** command to configure the VTP/CDP packet transparent transmission switch. After the switch is enabled, VTP/CDP packets are transparently transmitted based on the VLAN.
- Run the **rip tunnel** command to configure the RIP Layer 2 transparent transmission switch. After the transparent transmission switch is enabled, RIP packets can be transparently transmitted at Layer 2 based on VLAN on the device without running the RIP protocol.
- Run the **l3-protocol tunnel** command to configure the L2 transparent transmission for the L3 protocol packets except RIP and OSPF packets. After this function is enabled, the L3 protocol packets can be transparently transmitted at Layer 2 based on VLAN on the device without running the L3 protocol.
- Run the **ipv6 dad proxy** command to configure the DAD proxy (duplicate address detect proxy). DAD proxy prevents repeated LLA configuration on the user side.

Step 5 Configure the VLAN security function.

The security function is used to prevent malicious users from attacking the system by forging the IP address or MAC address of an authorized user. VMAC and anti-MAC spoofing conflict with S+C forwarding.

- Run the **security anti-ipspoofing** command to configure the anti-IPv4 spoofing function. After the anti-IPv4 spoofing function is enabled, the system automatically and dynamically binds the IPv4 address to the user. The packet can be transmitted upstream through the device only when the source IPv4 address of the packet is the same as the bound IPv4 address. Otherwise, the packet is discarded.
- Run the **security anti-ipv6spoofing** command to configure the anti-IPv6 spoofing function. After the anti-IPv6 spoofing function is enabled, the system automatically and dynamically binds the IPv6 address to the user. The packet can be transmitted upstream through the device only when the source IPv6 address of the packet is the same as the bound IPv6 address. Otherwise, the packet is discarded.
- Run the **security anti-macspoofing** command to configure the anti-MAC spoofing function. After the anti-MAC spoofing function is enabled, the system automatically and dynamically binds the MAC address to the traffic stream. When the source MAC address of the traffic stream is the same as the bound MAC address, the traffic stream can be upstream transmitted through the device. Otherwise, the packets are discarded.
- Run the **security arp-reply** command to enable the network-side ARP proxy response function. If the network-side ARP proxy response function is enabled, the system searches for user's going online information based on the destination IP address and VLAN after it receives network-side ARP request packets. If there is an online user, the system performs proxy response. If there is no online user, the system discards or forwards the ARP request packets based on the setting in the **security arp-reply unknown-policy** command. This prevents ARP request packets from being sent to user ports and reduces system resources.
- Run the **security ns-reply** command to enable the network-side NS proxy response function. If network-side NS proxy reply is enabled, the system searches for user's going online information based on the destination IP address and VLAN after it receives

network-side NS packets. If there is an online user, the system performs proxy response. If there is no online user, the system discards or forwards the NS packets based on the setting in the **security ns-reply unknown-policy** command. This prevents NS packets from being sent to user ports and reduces system resources.

- Run the **security bind-route-nd** command to configure the function of binding route with neighbor entry. After the function of binding route with neighbor entry is enabled, the system automatically generates the route and neighbor entry of a DHCPv6 user based on the user information recorded when the user goes online. This function reduces the effort of configuring static routes manually, prevents neighbor packets from being sent to the user side, and enhances system security.
- Run the **vmac** command to enable or disable VMAC. By default, VMAC is disabled.

Step 6 Commit to save the current parameters.

Run the **commit** command to commit the current parameter configuration of the VLAN service profile. After the configuration is completed, do run the **commit** command to make the configuration take effect.

Step 7 Bind the VLAN service profile to the VLAN.

Run the **vlan bind service-profile** command to bind the configured VLAN service profile to a specified VLAN. After the binding, the VLAN-level feature control switch is based on the configuration of the VLAN service profile. Independent configuration commands for VLAN-based features are no longer effective.

----End

Result

You can query the configuration of the VLAN service profile by the **display vlan service-profile** command.

After a VLAN service profile is bound to a VLAN, regarding the parameters whose **Committed** state is **NotConfig**, the configuration commands that are independent of the VLAN take effect; other parameters adopt the control parameters of the profile. Modifying the feature parameters relevant to the VLAN does not take effect.

Example

Add VLAN service profile 3 and bind it to VLAN 100. The profile parameters are planned as follows:

- VLAN forwarding mode VLAN+MAC address (vlan-mac)
- BPDU transparent transmission: enabled
- Unknown multicast packet: discarded

Adopt the default values for other parameters.

```
huawei(config)#vlan service-profile profile-id 3
huawei(config-vlan-srvprof-3)#forwarding vlan-mac
huawei(config-vlan-srvprof-3)#bpdu tunnel enable
huawei(config-vlan-srvprof-3)#packet-policy multicast discard
huawei(config-vlan-srvprof-3)#commit
huawei(config-vlan-srvprof-3)#quit
huawei(config)#vlan bind service-profile 100 profile-id 3
```

13.3.10 Reference Standards and Protocols

The following lists the reference standards and protocols of this feature:

Table 13-23 Reference standards and protocols of VLAN feature

Document name	Description
IEEE 802.1Q	IEEE standards for Local and metropolitan area networks-Virtual Bridged Local Area Networks
IEEE 802.1ad	Virtual Bridged Local Area Networks Amendment 4: Provider Bridges
RFC3069	VLAN Aggregation for Efficient IP Address Allocation

13.4 Service Flow

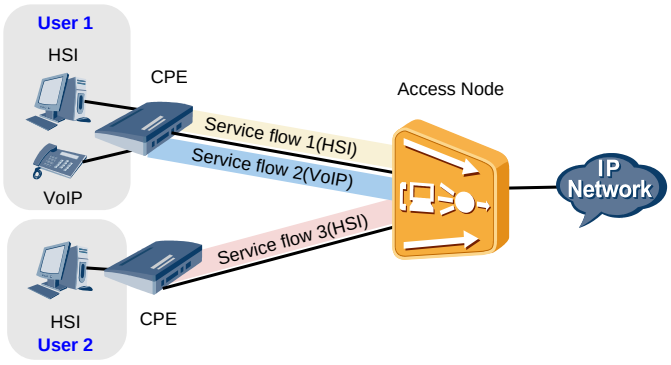
13.4.1 Introduction

An access device provides services to a large number of users, and each user requires multiple types of services (for example, HSI and VoIP services). The access device is required to differentiate between different user services when processing user packets, so that the services do not interfere with each other. To address this requirement, the access node provides the service flow feature.

Definition

Service flow, also called service port, is a result of traffic classification by characteristics of an Ethernet packet on a physical or logical port. Service flow is also a Layer 2 logical channel that carries services between the access node and users (Specify the Layer 2 forwarding path).

The following figure shows service flows model.



Services from different users (service flows 1 and 3) or services from the same user (service flows 1 and 2) are carried over different service flows.

Purpose

A service flow is the basis for provisioning each type of service on the access node. That is, the configuration of service flows is mandatory. In addition to traffic classification, a service flow is the smallest unit of user service processing. Hence, differentiated and fine-grained management, such as QoS, line identification, and security policies, of user services can be implemented based on service flows.

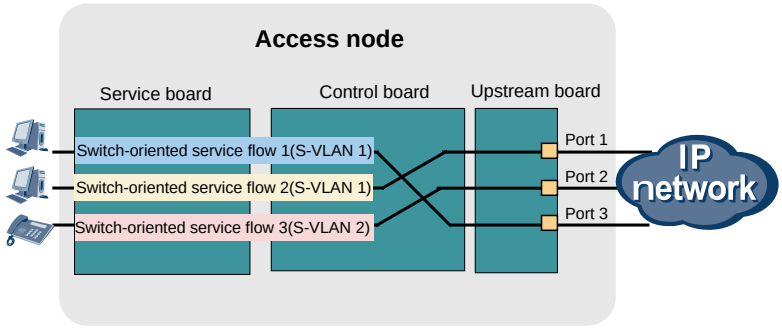
13.4.2 Principle

Service Flow Classification

Based on the service type, service flows can be classified as follows:

- **Switch-oriented service flow**: It is a Layer 2 service channel from the user access port on a device to the VLAN (S-VLAN) of the switching core. In an S-VLAN, one upstream port can be bound to multiple switch-oriented service flows and one switch-oriented service flow can be bound to multiple upstream ports. Since user information (such as VPI or VCI) is terminated at service boards. S-VLAN based switching is performed on the MA5600T/MA5603T/MA5608T. Therefore, the service flows transmitted in the Layer 2 service channel are called switch-oriented service flows.

  The following figure shows the schematic diagram for switch-oriented service flows.

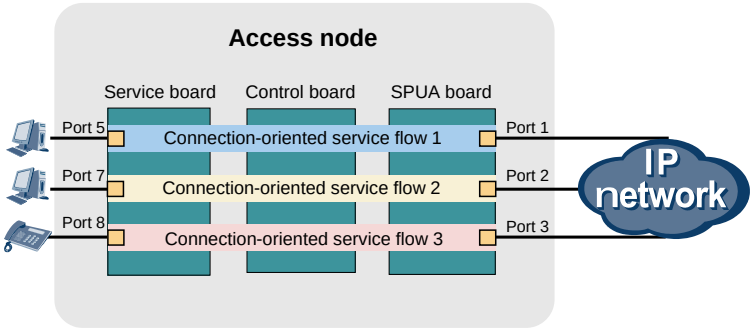


NOTE

Switching on the MA5600T/MA5603T/MA5608T: When a packet arrives at the MA5600T/MA5603T/MA5608T, the MA5600T/MA5603T/MA5608T selects an egress port and forwards the packet from this port. This process is called switching (or forwarding) on the MA5600T/MA5603T/MA5608T.

- **Connection-oriented service flow**: It is a point-to-point Layer 2 service channel from the user access or cascading port on a device to a network-side port. A connection-oriented service flow can be considered a transparent transmission tunnel between two ends (source and destination). A transparent channel is formed between the user side port and the network side port. Packets are transparently transmitted through this channel. Different ports can use the same VLAN without interfering with each other. On this channel, the MAC address needs not be learned.

  The following figure shows the schematic diagram for connection-oriented service flows.



- **ATM over Ethernet (AoE) emulation service flow**: It is a service channel that is created for ATM PWE3 emulation services. AoE emulation services refer to the traditional ATM services that are transmitted by access devices in an Ethernet network. In AoE emulation technology, ATM cells are not parsed and reassembled. Instead, they are considered the payload of PWE3 packets or Ethernet packets and are transparently transmitted across the network. AoE emulation service flows can be regarded as special switch-oriented service flows. The difference between the two types of service flows lies in that switch-oriented service flows use S-VLANs but AoE emulation service flows do not use S-VLANs.

All parameters for user services are specified for each service flow during the creation of the service flow. The values of these parameters determine the Layer 2 forwarding path and processing mode for the user services.

Traffic Classification Rules

Packets are classified based on traffic classification rules, with each type of packets corresponding to a service flow. The traffic classification rule for a service flow is specified when you run the **service-port** command to add the service flow. After packets are mapped to different service flows, the system processes the service flows based on the configured policies (including VLAN translation policy and QoS policy).

Traffic classification rules supported by the MA5600T/MA5603T/MA5608T can be generalized into two types:

- Classification based on physical ports or logical ports.
- Classification based on Ethernet packet attributes (such as priority). For details, see 14.5 Traffic Classification.

The following table lists the traffic classification rules supported by the MA5600T/MA5603T/MA5608T.

Table 13-24 Traffic classification rules supported by the MA5600T/MA5603T/MA5608T

Service Flow Type	Classification Rule Based on Physical Ports or Logical Ports		Classification Rule Based on Ethernet Packet Attributes
	Physical Port	Logical Port Identifier	
Switch-oriented service	Users can access the MA5600T/MA5603T/MA5608T through the following physical	• xDSL access mode: VC ports (PVCs) with VPI/VCI VC	Switch-oriented service flows support all the classification rules

Service Flow Type	Classification Rule Based on Physical Ports or Logical Ports		Classification Rule Based on Ethernet Packet Attributes
	Physical Port	Logical Port Identifier	
flow	ports: • ATM • xDSL • xPON • Ethernet access • Ethernet ports on ONTs in xPON access mode: The service flow created using this type of port is an end-to-end service flow between an OLT and an ONT.	ports can be automatically learned from the user side or be specified by users while a service flow is created. • GPON access mode: GEM port • xPON access mode: IPHOST voice port on an ONT. The IPHOST voice port is a virtual port of an ONT and implements communication between the voice chip and GMAC chip. The service flow created using this port is an end-to-end (voice) service flow between an OLT and an ONT.	Transparent LAN service (TLS) flows are also called other-all service flows.
Connection-oriented service flow	A destination port is usually a user-side port, cascading port, or user-side logical port. It can be ports on GPON, OPGD, OPGE, SPUA or SPUC boards. A source port is usually a network-side port. MA5600T/MA5603T/MA5608TOnly ports on the SPUA/SPUC board can be used as source ports.	Only the GEM port (a logical port) on the MA5600T/MA5603T/MA5608T can function as the destination port.	Only classification based on C-VLAN and classification based on double-tagged VLAN are supported. • Connection-oriented service flows between the GPON board and SPUA/SPUC board support only classification based on C-VLAN. • Connection-oriented service flows between the OPGD/OPGE/SPUA/SPUC board and SPUA/SPUC board support classification based on C-VLAN and

Service Flow Type	Classification Rule Based on Physical Ports or Logical Ports		Classification Rule Based on Ethernet Packet Attributes
	Physical Port	Logical Port Identifier	
			classification based on double-tagged VLAN.
AoE emulation service flow	Users can access the MA5600T/MA5603T/ MA5608T through the following physical ports: • ATM • ADSL • VDSL (only for the ATM access mode) • SHDSL (only for the ATM access mode)	VC ports (PVCs) with VPI/VCI. VC ports can be automatically learned from the user side or be specified by users while a service flow is created.	Supports only the single-service classification rule.

Attributes of Service Flows

Table 13-25 Attributes of service flows

Name	Description	Configuration Command
Service flow index	Each service flow has a unique index on the MA5600T/MA5603T/MA5608T. The index can be used when other features based on the service flow are configured. For example, you can configure a user as a multicast user simply by specifying the service flow index.	**service-port** *index*
Description of the service flow	There are usually a large number of service flows in an access node. The description (for example, the area in which the user is located and service characteristics) of a service flow facilitates service maintenance.	**service-port desc**
Description of the remote device of the service flow	The description includes information about the remote device of the user side of a service flow and helps determine the terminal type when many types of terminals are used, which facilitates maintenance.	**service-port remote-desc**
Administrative status	Users can pause and resume the transmission of services on a service flow by setting the administrative status of the service flow.	**service-port adminstatus**

Automatic Service Flow Creation

The OLT supports pre-configuration of service flow creation policies on a PON port. Then, after an ONU goes online, services flows are automatically created based on the pre-configured policies. This function simplifies the configuration process and improves the installation and deployment efficiency.

O&M for these automatically created service flows is the same as that for commonly created service flows.

Figure 13-16 Process of automatic service flow creation (no ONU is manually added)

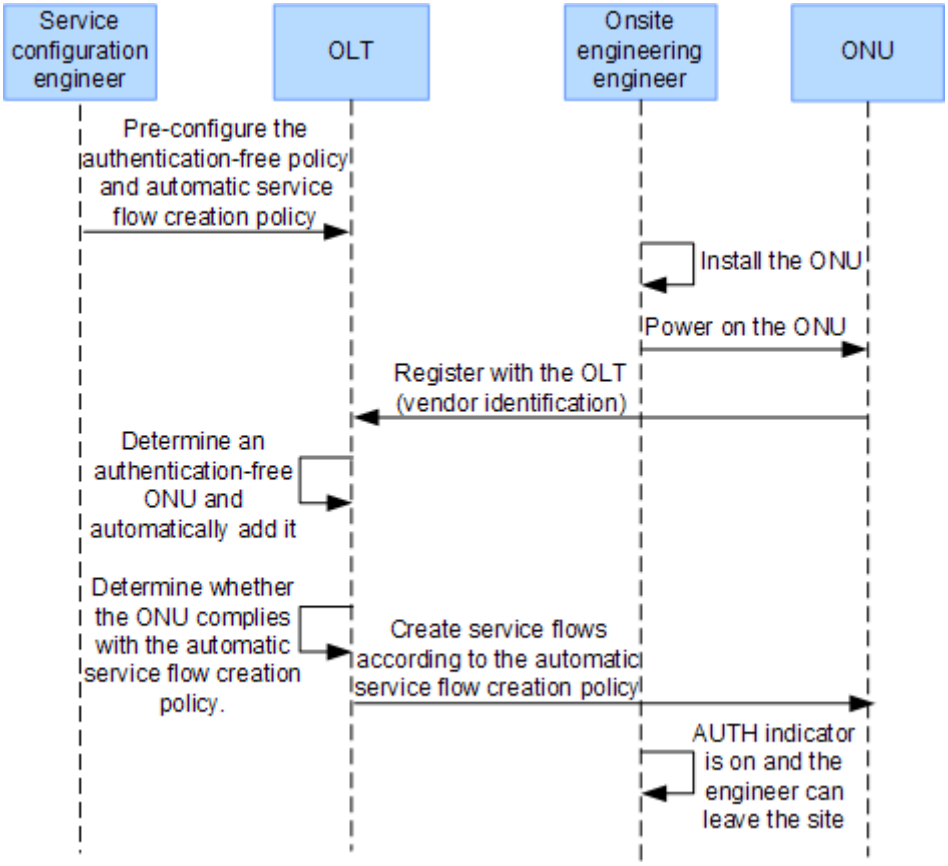
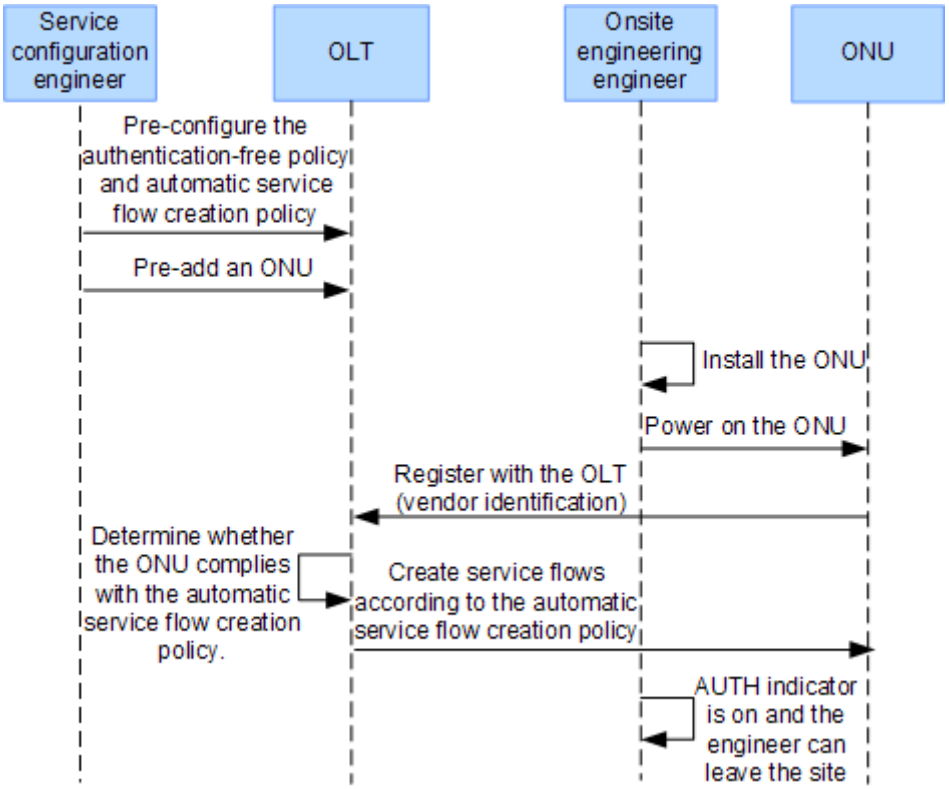


Figure 13-17 Process of automatic service flow creation (ONU is manually added)



13.4.3 Configuration

The method of configuring service flows varies according to the access mode. You can locate the desired configuration description according to the access mode.

Creating an xDSL or Ethernet Service Flow

An xDSL or Ethernet service flow is a service channel connecting the user side to the network side. Configure a service flow before provisioning services.

Context

A service flow can carry a single service or multiple services. A multi-service flow is used to carry triple-play services. A service port can carry a single service flow or multiple service flows. If a service port carries multiple service flows, the OLT supports traffic classification and identifies users or services based on:

- User-side VLAN.
- User-side service encapsulation type. For example, PPPoE is used for the Internet access service and IPoE is used for multicast services.
- VLAN+user-side packet priority. For example, the priority of multicast services is higher than that of the Internet access service.
- VLAN+user-side service encapsulation type.

Before creating a service flow, run the **display traffic table** command to query whether the desired service flow has existed in the system. If no desired service flow is available, run the

traffic table ip command to create a traffic profile for the service flow. The OLT provides seven default traffic profiles with IDs ranging from 0 to 6.

Table 13-26 lists the default settings of a service flow.

Table 13-26 Default settings of a service flow

Parameter	Default Setting
Traffic profile ID	0-6
Management status	Activated

Procedure

Create a service flow or multiple service flows in batches.

- Run the **service-port** command to create a service flow.
  - Single-service flow:

    Select **single-service** or do not set **multi-service**.
  - Multi-service flow based on the user-side VLAN:

    Select **multi-service user-vlan** { **untagged** | *user-vlanid* | **priority-tagged** | **other-all** }.
    - **untagged**: If this parameter is specified, user packets do not carry a tag.
    - *user-vlanid*: If this parameter is specified, user packets carry a tag, which is the customer VLAN (C-VLAN).
    - **priority-tagged**: If this parameter is specified, the priorities of user packets range from 0 to 7. (The highest priority is 7.)
    - **other-all**: If this parameter is specified, the created service flow carries QinQ transparent LAN service (TLS) services for enterprises. User packets are matched based on the specified user VLAN (or untagged attribute). The unmatched packets are transmitted on the TLS service flow to the upper-layer network.
  - Multi-service flow based on the user-side service encapsulation mode:

    Select **multi-service user-encap** *user-encap*.
  - Multi-service flow based on the VLAN+user-side packet priority (802.1p)

    Select **multi-service user-8021p** *user-8021p* [**user-vlan** *user-vlanid*].
  - Multi-service flow based on the VLAN+user-side service encapsulation mode (user-encap)

    Select **multi-service user-vlan** { **untagged** | *user-vlanid* | **priority-tagged** } **user-encap** *user-encap*.

NOTE

- The OLT supports the service flow configuration by index. Each service flow has a unique index. In this manner, users do not need to a large number of flow parameters, thereby simplifying service flow configurations. When creating a service flow, the service flow index parameter *index* is optional. If *index* is not set, the OLT automatically allocates an idle index starting from the configured maximum index, regardless of whether the maximum index has been deleted. If the idle index is greater than the upper index threshold, the OLT searches for the new idle index starting from 0.
- **vlan** is an S-VLAN, which can only be a smart or MUX VLAN.

- The access mode can be ATM or PTM. In ATM access mode, the VPI and VCI must be input and the same as the VPI and VCI of the access terminal.
- **rx-cttr** is the same as **outbound** in terms of meaning and function. Either parameter indicates the index of the service flow from the network side to the user side. **tx-cttr** is the same as **inbound** in terms of meaning and function. Either parameter indicates the index of the service flow from the user side to the network side.

- Run the **multi-service-port** command to create service flows in batches.

Step 1 (Optional) Run the **service-port desc** command to configure the description of the service flow. The description includes the purpose of the service flow creation and the services carried over the service flow. Therefore, the description of the service flow facilitates maintenance.

Step 2 (Optional) Run the **service-port *index* adminstatus** command to configure the management status of the service flow. A service port activated by default.

A service can be provisioned at two levels: port level and service flow level. The service takes effect only after both the access port and service flow are activated.

----End

Example

The following configurations are used as an example to provision the Internet access service for a home user:

- The planned configurations of the user are as follows:
  - The access mode is ADSL2+.
  - The rate of the Internet access service is 3072 Kbit/s.
  - The user connects to Ethernet port 1 on the ONT.
- No proper traffic profile is available.
- The user has not registered. Do not provision the Internet access service for the user.

```
huawei(config)#display traffic table ip from-index 0
{ <cr>|to-index<K> }:

 Command:
       display traffic table ip from-index 0
 ------------------------------------------------------------------------------
  TID CIR     CBS     PIR     PBS     Pri Copy-policy       Pri-Policy
     (kbps)   (bytes)  (kbps)   (bytes)
 ------------------------------------------------------------------------------
    0 1024    34768    2048    69536     6 -                  tag-pri
    1 2496    81872    4992    163744    6 -                  tag-pri
    2 512     18384    1024    36768     0 -                  tag-pri
    3 576     20432    1152    40864     2 -                  tag-pri
    4 64      4048     128     8096      4 -                  tag-pri
    5 2048    67536    4096    135072    0 -                  tag-pri
    6 off     off      off     off       0 -                  tag-pri
 ------------------------------------------------------------------------------
 Total Num : 7
huawei(config)#traffic table ip index 8 cir 3072 priority 4 priority-policy loca
l-Setting
 Create traffic descriptor record successfully
 ------------------------------------------------
 TD Index           : 8
```

```
  TD Name           : ip-traffic-table_8
  Priority          : 4
  Copy Priority      : -
  Mapping Index      : -
  CTAG Mapping Priority: -
  CTAG Mapping Index  : -
  CTAG Default Priority: 0
  Priority Policy    : local-pri
  CIR               : 3072 kbps
  CBS               : 100304 bytes
  PIR               : 6144 kbps
  PBS               : 198608 bytes
  Color policy       : dei
  Referenced Status   : not used
  ------------------------------------------------
huawei(config)#service-port 3 vlan 100 adsl 0/2/1 vpi 1 vci 35 inbound traffic-table
 index 8 outbound traffic-table index 8
huawei(config)#service-port 3 adminstatus disable
```

The following configurations are used as an example to provision the Internet access service for a home user:

- The planned configurations of the user are as follows:
  - The access mode is ADSL2+.
  - The rate of the Internet access service is 2048 Kbit/s.
  - The Internet access service is in multi-service mode to facilitate follow-up service expansion.
  - The user is identified by C-VLAN.
  - The service and physical location of the user is identified by S-VLAN.
  - The ID of the C-VLAN is 10.
  - The ID of the S-VLAN is 50.
- No proper traffic profile is available.
- The description of the service flow is added to facilitate maintenance.
- The Internet access service needs to be provisioned immediately.

```
huawei(config)#display traffic table ip from-index 0
{ <cr>|to-index<K> }:

  Command:
        display traffic table ip from-index 0
  ------------------------------------------------------------------------------
   TID CIR     CBS     PIR     PBS     Pri Copy-policy       Pri-Policy
       (kbps)  (bytes) (kbps)  (bytes)
  ------------------------------------------------------------------------------
     0 1024    34768   2048    69536     6 -                  tag-pri
     1 2496    81872   4992    163744    6 -                  tag-pri
     2 512     18384   1024    36768     0 -                  tag-pri
     3 576     20432   1152    40864     2 -                  tag-pri
     4 64      4048    128     8096      4 -                  tag-pri
     5 2048    67536   4096    135072    0 -                  tag-pri
     6 off     off     off     off       0 -                  tag-pri
  ------------------------------------------------------------------------------
  Total Num : 7
```

```
huawei(config)#service-port 4 vlan 50 adsl 0/2/1 vpi 1 vci 39 multi-service
user-vlan 10 inbound traffic-table index 5 outbound traffic-table index 5
huawei(config)#service-port desc 4 description HW_adsl/VlanID:50/uservlan/10
```

The following configurations are used as an example to provision the Internet access service for a commercial user:

- The planned configurations of the user are as follows:
  - The access mode is VDSL.
  - The rate of the Internet access service is 8192 Kbit/s.
  - The Internet access service is in multi-service mode to facilitate follow-up service expansion.
  - The user is identified by C-VLAN.
  - The service and physical location of the user is identified by S-VLAN.
  - The ID of the C-VLAN is 10.
  - The ID of the S-VLAN is 50.
- No proper traffic profile is available.
- The description of the service flow is added to facilitate maintenance.
- The Internet access service needs to be provisioned immediately.

```
huawei(config)#display traffic table ip from-index 0
{ <cr>|to-index<K> }:

 Command:
        display traffic table ip from-index 0
 ------------------------------------------------------------------------------
  TID CIR     CBS     PIR     PBS     Pri Copy-policy       Pri-Policy
      (kbps)  (bytes) (kbps)  (bytes)
 ------------------------------------------------------------------------------
    0 1024    34768   2048    69536     6 -                  tag-pri
    1 2496    81872   4992    163744    6 -                  tag-pri
    2 512     18384   1024    36768     0 -                  tag-pri
    3 576     20432   1152    40864     2 -                  tag-pri
    4 64      4048    128     8096      4 -                  tag-pri
    5 2048    67536   4096    135072    0 -                  tag-pri
    6 off     off     off     off       0 -                  tag-pri
    8 3072    100304  6144    198608    4 -                  local-pri
 ------------------------------------------------------------------------------
 Total Num : 8
huawei(config)#traffic table ip index 9 cir 8192 priority 4 priority-policy
 local-Setting
 Create traffic descriptor record successfully
 ------------------------------------------------
 TD Index           : 9
 TD Name            : ip-traffic-table_9
 Priority           : 4
 Copy Priority       : -
 Mapping Index       : -
 CTAG Mapping Priority: -
 CTAG Mapping Index   : -
 CTAG Default Priority: 0
 Priority Policy     : local-pri
 CIR               : 8192 kbps
```

```
  CBS               : 264144 bytes
  PIR               : 16384 kbps
  PBS               : 526288 bytes
  Color policy       : dei
  Referenced Status   : not used
  ------------------------------------------------
huawei(config)#service-port 6 vlan 50 vdsl mode ptm 0/4/1 multi-service user-vlan
 10 inbound traffic-table index 9 outbound traffic-table index 9
huawei(config)#service-port desc 6 description HW_vdsl/VlanID:50/uservlan:10
```

The following configurations are used as an example to provision the Internet access service:

- The access port is 0/4/1.
- The ID of the C-VLAN is 2.
- The ID of the upstream VLAN is 100.
- Service profile 8 named **huawei** is used.
- The committed information rate (CIR) is 10240.
- The upstream priority is 0.
- The packet priority policy is determined by the traffic profile for local priority scheduling.

```
huawei(config)#traffic table ip index 8 name net cir 10240 priority 0 priority-policy
 local-setting
  Create traffic descriptor record successfully
  ------------------------------------------------
  TD Index           : 8
  TD Name            : huawei
  Priority           : 0
  Copy Priority       : -
  CTAG Mapping Priority: -
  CTAG Default Priority: 0
  Priority Policy     : local-pri
  CIR               : 10240 kbps
  CBS               : 329680 bytes
  PIR               : 20480 kbps
  PBS               : 657360 bytes
  Color Mode         : color-blind
  Referenced Status   : not used
  ------------------------------------------------
huawei(config)#service-port 1 vlan 100 eth 0/4/1 multi-service user-vlan 2 rx-cttr 8
tx-cttr 8
```

Creating a GPON Service Flow (in Distributed Mode)

A GPON service flow is a service channel connecting the user side to the network side in distributed mode. Configure a service flow before provisioning services.

Context

A service flow can carry a single service or multiple services. A multi-service flow is used to carry triple-play services. A service port can carry a single service flow or multiple service flows. If a service port carries multiple service flows, the OLT supports traffic classification and identifies users or services based on:

- User-side VLAN.
- User-side service encapsulation type.
- VLAN+user-side packet priority.
- VLAN+user-side service encapsulation type.

Table 13-27 lists the default settings of a service flow.

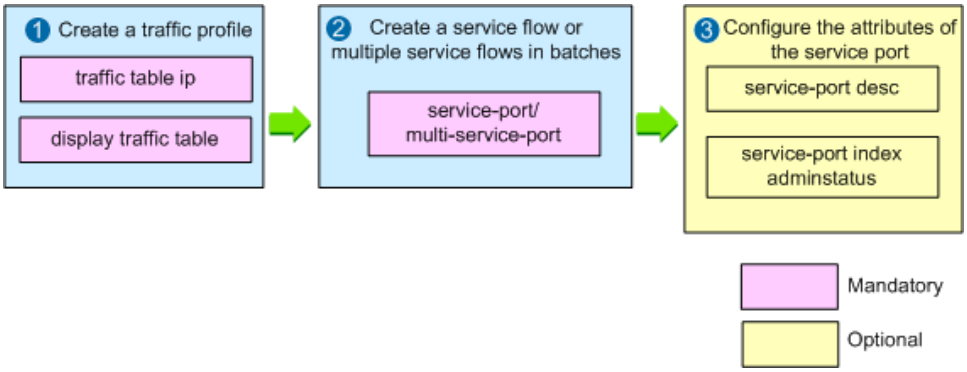
Table 13-27 Default settings of a service flow

Parameter	Default Setting
Traffic profile ID	0-6
Management status	Activated

Configuration Process

Figure 13-18 shows the process of creating a GPON service flow.

Figure 13-18 Process of creating a GPON service flow



Procedure

Create a traffic profile.

Run the **traffic table ip** command to create a traffic profile. The OLT provides seven default traffic profiles with IDs ranging from 0 to 6.

Before creating a service flow, run the **display traffic table** command to query whether the desired service flow has existed in the system. If no desired service flow is available, run the **traffic table ip** command to create a traffic profile for the service flow. For details about traffic profiles, see Configuring Rate Limitation Based on Service Port.

Step 1 Create a service flow or multiple service flows in batches.

- Run the **service-port** command to create a service flow.
  - Single-service flow:

    Do not set **multi-service**.
  - Multi-service flow based on the user-side VLAN:

Select **multi-service user-vlan** { **untagged** | *user-vlanid* | **priority-tagged** | **other-all** }.

- **untagged**: If this parameter is specified, user packets do not carry a tag.
- *user-vlanid*: If this parameter is specified, user packets carry a tag, which is the customer VLAN (C-VLAN).
- **priority-tagged**: If this parameter is specified, the priorities of user packets range from 0 to 7. (The highest priority is 7.)
- **other-all**: If this parameter is specified, the created service flow carries QinQ transparent LAN service (TLS) services for enterprises. User packets are matched based on the specified user VLAN (or untagged attribute). The unmatched packets are transmitted on the TLS service flow to the upper-layer network.

– Multi-service flow based on the user-side service encapsulation mode:

Select **multi-service user-encap** *user-encap*.

– Multi-service flow based on the VLAN+user-side packet priority (802.1p)

Select **multi-service user-8021p** *user-8021p* [**user-vlan** *user-vlanid*].

– Multi-service flow based on the VLAN+user-side service encapsulation mode (user-encap)

Select **multi-service user-vlan** { **untagged** | *user-vlanid* | **priority-tagged** } **user-encap** *user-encap*.

NOTE

- The OLT supports the service flow configuration by index. Each service flow has a unique index. In this manner, users do not need to a large number of flow parameters, thereby simplifying service flow configurations. When creating a service flow, the service flow index parameter *index* is optional. If *index* is not set, the OLT automatically allocates an idle index starting from the minimum unused value.
- **vlan** is an S-VLAN, which can only be a smart or MUX VLAN.
- **rx-cttr** is the same as **outbound** in terms of meaning and function. Either parameter indicates the index of the service flow from the network side to the user side. **tx-cttr** is the same as **inbound** in terms of meaning and function. Either parameter indicates the index of the service flow from the user side to the network side. The traffic profile bound to the service flow is the one created in Step 1.

- Run the **multi-service-port** command to create service flows in batches.

Step 2 Configure the attributes of the service port. Configure the attributes of the service port according to requirements.

- Run the **service-port desc** command to configure the description of the service flow. The description includes the purpose of the service flow creation and the services carried over the service flow. Therefore, the description of the service flow facilitates maintenance.
- Run the **service-port *index* adminstatus** command to configure the management status of the service flow. A service port activated by default.

  A service can be provisioned at two levels: port level and service flow level. The service takes effect only after both the access port and service flow are activated.

----End

Example

The following configurations are used as an example to provision the Internet access service for a user:

- GPON port 0/2/0 on an OLT connects to an ONT.
- The planned configurations of the user are as follows:
  - The rate of the Internet access service is 4096 Kbit/s.
  - The index of the GEM port carrying the Internet access service is 135.
  - The ID of the service VLAN (S-VLAN) is 1000.
- No proper traffic profile is available and traffic profile 10 is created.
- The user has not registered. Do not provision the Internet access service for the user.

```
huawei(config)#traffic table ip index 10 cir 4096 priority 3 priority-policy loc
al-Setting
  Create traffic descriptor record successfully
  -------------------------------------------------
  TD Index           : 10
  TD Name            : ip-traffic-table_10
  Priority           : 3
  Copy Priority       : -
  Mapping Index       : -
  CTAG Mapping Priority: -
  CTAG Mapping Index   : -
  CTAG Default Priority: 0
  Priority Policy     : local-pri
  CIR                : 4096 kbps
  CBS                : 133072 bytes
  PIR                : 8192 kbps
  PBS                : 264144 bytes
  Fix                : 0 kbps
  CAR Threshold Profile: -
  Color Mode         : color-blind
  Color policy        : dei
  Referenced Status   : not used
  -------------------------------------------------
huawei(config)#service-port 5 vlan 1000 gpon 0/2/0 gemport 135 inbound
 traffic-table index 10 outbound traffic-table index 10
huawei(config)#service-port 5 adminstatus disable
```

The following configurations are used as an example to provision the Internet access service for a commercial user:

- GPON port 0/2/0 on an OLT connects to an ONT.
- The planned configurations of the user are as follows:
  - The rate of the Internet access service is 8192 Kbit/s.
  - The Internet access service is in multi-service mode to facilitate follow-up service expansion.
  - The user is identified by C-VLAN.
  - The ID of the S-VLAN is 1023.
  - The ID of the C-VLAN is 100.
  - The index of the GEM port carrying the Internet access service is 130.
- No proper traffic profile is available and traffic profile 8 is created.
- The description of the service flow is added to facilitate maintenance.
- The Internet access service needs to be provisioned immediately.

```
huawei(config)#traffic table ip index 8 cir 8192 priority 4 priority-policy loca
l-Setting
  Create traffic descriptor record successfully
  ------------------------------------------------
  TD Index           : 8
  TD Name            : ip-traffic-table_8
  Priority           : 4
  Copy Priority       : -
  Mapping Index       : -
  CTAG Mapping Priority: -
  CTAG Mapping Index   : -
  CTAG Default Priority: 0
  Priority Policy     : local-pri
  CIR                : 8192 kbps
  CBS                : 264144 bytes
  PIR                : 16384 kbps
  PBS                : 526288 bytes
  Fix                : 0 kbps
  CAR Threshold Profile: -
  Color Mode         : color-blind
  Color policy        : dei
  Referenced Status   : not used
  ------------------------------------------------
huawei(config)#service-port 10 vlan 1023 gpon 0/2/0 gemport 130 multi-service
user-vlan 100 inbound traffic-table index 8 outbound traffic-table index 8
huawei(config)#service-port desc 10 description gpon/vlanid:1023/uservlan:100
```

Creating a GPON Service Flow (in Profile Mode with Universal Configurations)

A GPON service flow (with universal configurations) is a service channel connecting the user side to the network side in profile mode. Configure a service flow before provisioning services.

Context

A service flow can carry a single service or multiple services. A multi-service flow is used to carry triple-play services. A service port can carry a single service flow or multiple service flows. If a service port carries multiple service flows, the OLT supports traffic classification and identifies users or services based on:

- User-side VLAN.
- User-side service encapsulation type.
- VLAN+user-side packet priority.
- VLAN+user-side service encapsulation type.

Table 13-28 lists the default settings of a service flow.

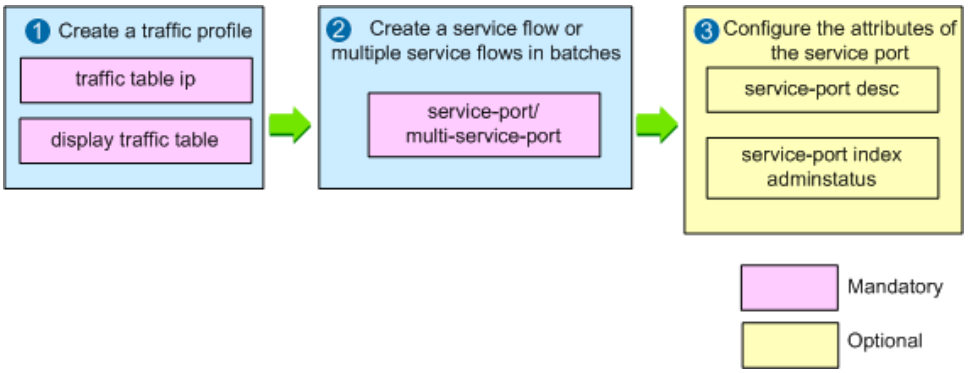
Table 13-28 Default settings of a service flow

Parameter	Default Setting
Traffic profile ID	0-6
Management status	Activated

Configuration Process

Figure 13-19 shows the process of creating a GPON service flow.

Figure 13-19 Process of creating a GPON service flow



Procedure

Create a traffic profile.

Run the **traffic table ip** command to create a traffic profile. The OLT provides seven default traffic profiles with IDs ranging from 0 to 6.

Before creating a service flow, run the **display traffic table** command to query whether the desired service flow has existed in the system. If no desired service flow is available, run the **traffic table ip** command to create a traffic profile for the service flow. For details about traffic profiles, see Configuring Rate Limitation Based on Service Port.

Step 1 Create a service flow or multiple service flows in batches.

- Run the **service-port** command to create a service flow.
  - Single-service flow:

    Do not set **multi-service**.
  - Multi-service flow based on the user-side VLAN:

    Select **multi-service user-vlan** { **untagged** | *user-vlanid* | **priority-tagged** | **other-all** }.
    - **untagged**: If this parameter is specified, user packets do not carry a tag.
    - *user-vlanid*: If this parameter is specified, user packets carry a tag, which is the customer VLAN (C-VLAN).
    - **priority-tagged**: If this parameter is specified, the priorities of user packets range from 0 to 7. (The highest priority is 7.)
    - **other-all**: If this parameter is specified, the created service flow carries QinQ transparent LAN service (TLS) services for enterprises. User packets are matched based on the specified user VLAN (or untagged attribute). The unmatched packets are transmitted on the TLS service flow to the upper-layer network.
  - Multi-service flow based on the user-side service encapsulation mode:

Select **multi-service user-encap** *user-encap*.

- Multi-service flow based on the VLAN+user-side packet priority (802.1p)

  Select **multi-service user-8021p** *user-8021p* [**user-vlan** *user-vlanid*].

- Multi-service flow based on the VLAN+user-side service encapsulation mode (user-encap)

  Select **multi-service user-vlan** { **untagged** | *user-vlanid* | **priority-tagged** } **user-encap** *user-encap*.

NOTE

- The OLT supports the service flow configuration by index. Each service flow has a unique index. In this manner, users do not need to a large number of flow parameters, thereby simplifying service flow configurations. When creating a service flow, the service flow index parameter *index* is optional. If *index* is not set, the OLT automatically allocates an idle index starting from the configured maximum index, regardless of whether the maximum index has been deleted. If the idle index is greater than the upper index threshold, the OLT searches for the new idle index starting from 0.
- **vlan** is an S-VLAN, which can only be a smart or MUX VLAN.
- **rx-cttr** is the same as **outbound** in terms of meaning and function. Either parameter indicates the index of the service flow from the network side to the user side. **tx-cttr** is the same as **inbound** in terms of meaning and function. Either parameter indicates the index of the service flow from the user side to the network side. The traffic profile bound to the service flow is the one created in Step 1.

- Run the **multi-service-port** command to create service flows in batches.

Step 2 Configure the attributes of the service port. Configure the attributes of the service port according to requirements.

- Run the **service-port desc** command to configure the description of the service flow. The description includes the purpose of the service flow creation and the services carried over the service flow. Therefore, the description of the service flow facilitates maintenance.
- Run the **service-port *index* adminstatus** command to configure the management status of the service flow. A service port activated by default.

  A service can be provisioned at two levels: port level and service flow level. The service takes effect only after both the access port and service flow are activated.

----End

Example

The following configurations are used as an example to provision the Internet access service for a user:

- GPON port 0/2/0 on an OLT connects to ONT 1.
- The planned configurations of the user are as follows:
  - The rate of the Internet access service is 4096 Kbit/s.
  - The index of the GEM port carrying the Internet access service is 126.
  - The ID of the service VLAN (S-VLAN) is 1000.
- No proper traffic profile is available and traffic profile 10 is created.
- The user has not registered. Do not provision the Internet access service for the user.

```
huawei(config)#traffic table ip index 10 cir 4096 priority 3 priority-policy local-Setting
  Create traffic descriptor record successfully
```

```
  ------------------------------------------------
  TD Index           : 10
  TD Name            : ip-traffic-table_10
  Priority           : 3
  Mapping Priority    : -
  Mapping Index       : -
  CTAG Mapping Priority: -
  CTAG Mapping Index   : -
  CTAG Default Priority: 0
  Priority Policy     : local-pri
  CIR                : 4096 kbps
  CBS                : 133072 bytes
  PIR                : 8192 kbps
  PBS                : 264144 bytes
  Fix                : 0 kbps
  CAR Threshold Profile: -
  Color Mode          : color-blind
  Color policy        : dei
  Referenced Status    : not used
  ------------------------------------------------
huawei(config)#service-port 5 vlan 1000 gpon OLT ont 1 gemport 126 inbound
traffic-table index 10 outbound traffic-table index 10
huawei(config)#service-port 5 adminstatus disable
```

The following configurations are used as an example to provision the Internet access service for a commercial user:

- GPON port 0/2/0 on an OLT connects to ONT 2.
- The planned configurations of the user are as follows:
  - The rate of the Internet access service is 8192 Kbit/s.
  - The Internet access service is in multi-service mode to facilitate follow-up service expansion.
  - The user is identified by C-VLAN.
  - The ID of the S-VLAN is 1023.
  - The ID of the C-VLAN is 100.
  - The index of the GEM port carrying the Internet access service is 126.
- No proper traffic profile is available and traffic profile 8 is created.
- The description of the service flow is added to facilitate maintenance.
- The Internet access service needs to be provisioned immediately.

```
huawei(config)#display traffic table ip from-index 0
{ <cr>|to-index<K> }:

  Command:
         display traffic table ip from-index 0
  ------------------------------------------------------------------------------
   TID CIR     CBS     PIR     PBS     Pri Copy-policy       Pri-Policy
       (kbps)  (bytes) (kbps)  (bytes)
  ------------------------------------------------------------------------------
     0 1024    34768   2048    69536     6 -                  tag-pri
     1 2496    81872   4992    163744    6 -                  tag-pri
     2 512     18384   1024    36768     0 -                  tag-pri
     3 576     20432   1152    40864     2 -                  tag-pri
```

```
    4 64      4048    128     8096      4 -                  tag-pri
    5 2048    67536   4096    135072    0 -                   tag-pri
    6 off     off     off     off       0 -                  tag-pri
  ------------------------------------------------------------------------------
  Total Num : 7
huawei(config)#traffic table ip index 8 cir 8192 priority 4 priority-policy
 local-Setting
  Create traffic descriptor record successfully
  ------------------------------------------------
  TD Index           : 8
  TD Name            : ip-traffic-table_8
  Priority           : 4
  Copy Priority       : -
  Mapping Index       : -
  CTAG Mapping Priority: -
  CTAG Mapping Index  : -
  CTAG Default Priority: 0
  Priority Policy     : local-pri
  CIR               : 8192 kbps
  CBS               : 264144 bytes
  PIR               : 16384 kbps
  PBS               : 526288 bytes
  Fix               : 0 kbps
  CAR Threshold Profile: -
  Color Mode         : color-blind
  Color policy        : dei
  Referenced Status   : not used
  ------------------------------------------------
huawei(config)#service-port 10 vlan 1023 gpon 0/2/0 ont 2 gemport 126 multi-service
 user-vlan 100 inbound traffic-table index 8 outbound traffic-table index 8
huawei(config)#service-port desc 10 description gpon/Vlanid:1023/uservlan:100
```

Creating a GPON Service Flow in Profile Mode with Simplified Configurations

A GPON service flow in profile mode with simplified configurations is a service channel connecting the user side to the network side. The service flow with simplified configurations involves only the switching between a service VLAN (S-VLAN) and a customer VLAN (C-VLAN). Configure a service flow before provisioning services.

Context

A service port can carry a single service or multiple services. When a service port carries multiple services, the OLT supports traffic classification. In simplified configuration mode, the OLT supports only C-VLAN-based traffic classification.

Table 13-29 lists the default settings of a service flow.

Table 13-29 Default settings of a service flow

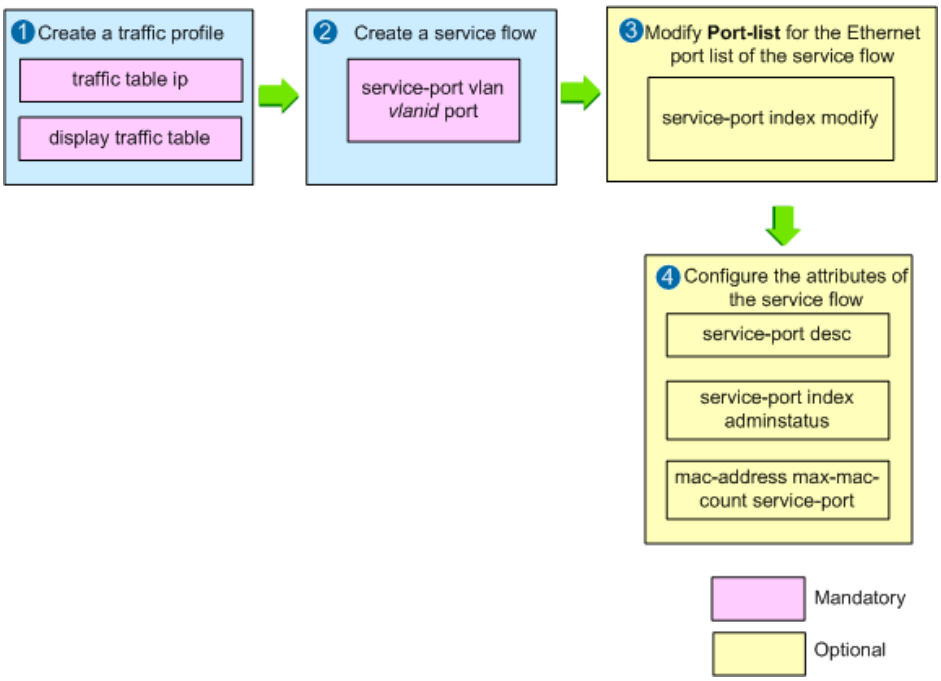
Parameter	Default Setting
Traffic profile ID	0-6
Management status	Activated
Maximum number of learnable MAC	1023

Parameter	Default Setting
addresses	

Configuration Process

Figure 13-20 shows the process of creating a GPON service flow.

Figure 13-20 Process of creating a GPON service flow



Procedure

Run the **traffic table ip** command to create a traffic profile.

The OLT provides seven default traffic profiles with IDs ranging from 0 to 6.

Before creating a service flow, run the **display traffic table** command to check whether a service flow meeting service requirements is available. If no such a service flow is available, create a traffic profile. For details about traffic profiles, see Configuring Rate Limitation Based on Service Port.

Step 1 Run the **service-port vlan *vlanid* port** command to create a service flow.

In simplified configuration mode, a service flow is an end-to-end configuration to a port on an optical network terminal (ONT) from the OLT. Then, users only need to pay attention to the switching between the S-VLAN and C-VLAN.

1. Select the type of an ONT port from **ont** *ontid* { **eth** | **iphost** }.
   - **eth**: an Ethernet port, which is used for the Internet access or multicast service.
   - **iphost**: a POTS port, which is used for the voice service.

2. Select a C-VLAN tag from **multi-service user-vlan** { **untagged** | *user-vlanid* }.
   - **untagged**: indicates that the user-side packets are untagged.
   - *user-vlanid*: indicates that the user-side packets are tagged. The tag value is the same as the C-VLAN tag.

The preceding parameters are set based on the ONT type. Therefore, specify the ONT type before creating a service flow.

- For a bridging ONT (SFU), the ONT port provisioning the POTS service is set to **iphost** and ONT ports provisioning other services are set to **eth**.
- For a gateway ONT (HGU), the ONT port provisioning the Internet access service with untagged packets is set to **iphost**; other services cannot be configured on physical ports but on the ONT.

Table 13-30 lists the commands involved in parameter settings.

Table 13-30 Commands involved in parameter settings

Service Type	Bridging ONT	Gateway ONT
Tagged high-speed Internet (HSI) service	**service-port** [*index*] **vlan** *vlanid* **port** *frameid/slotid/portid* **ont** *ontid* **eth** *port-index-list* **multi-service**...	**service-port** [*index*] **vlan** *vlanid* **port** *frameid/slotid/portid* **ont** *ontid* **multi-service**...
Untagged HSI service	**service-port** [*index*] **vlan** *vlanid* **port** *frameid/slotid/portid* **ont** *ontid* **eth** *port-index-list* **multi-service**...	**service-port** [*index*] **vlan** *vlanid* **port** *frameid/slotid/portid* **ont** *ontid* **iphost multi-service**... **NOTE** Services can be identified only after a native VLAN is configured for ONT ports.
Voice service	**service-port** [*index*] **vlan** *vlanid* **port** *frameid/slotid/portid* **ont** *ontid* **iphost multi-service**...	**service-port** [*index*] **vlan** *vlanid* **port** *frameid/slotid/portid* **ont** *ontid* **multi-service**...
IPTV service	**service-port** [*index*] **vlan** *vlanid* **port** *frameid/slotid/portid* **ont** *ontid* **eth** *port-index-list* **multi-service**...	**service-port** [*index*] **vlan** *vlanid* **port** *frameid/slotid/portid* **ont** *ontid* **multi-service**...

In the preceding table, "..." indicates the omitted command format.

NOTE

- The OLT supports the service flow configuration by index. Each service flow has a unique index. In this manner, users do not need to a large number of flow parameters, thereby simplifying service flow configurations. When creating a service flow, the service flow index parameter *index* is optional. If *index* is not set, the OLT automatically allocates an idle index starting from the configured maximum index, regardless of whether the maximum index has been deleted. If the idle index is greater than the upper index threshold, the OLT searches for the new idle index starting from 0.
- **vlan** is an S-VLAN, which can only be a smart or MUX VLAN.

- **rx-cttr** is the same as **outbound** in terms of meaning and function. Either parameter indicates the index of the service flow from the network side to the user side. **tx-cttr** is the same as **inbound** in terms of meaning and function. Either parameter indicates the index of the service flow from the user side to the network side. The traffic profile bound to the service flow is the one created in Step 1.

Step 2 (Optional) Run the **service-port *index* modify** command to modify **Port-list** for the Ethernet port list of the service flow.

Perform this step if the service planning for a user port is changed.

Step 3 (Optional) Configure the attributes of the service flow.

- Run the **service-port desc** command to configure the description of the service flow. The description includes the purpose of the service flow creation and the services carried over the service flow. Therefore, the description of the service flow facilitates maintenance.
- Run the **service-port *index* adminstatus** command to configure the management status of the service flow. A service port activated by default.

  A service can be provisioned at two levels: port level and service flow level. The service takes effect only after both the access port and service flow are activated.
- Run the **mac-address max-mac-count service-port** command to set the maximum number of MAC addresses learned by the service flow. This configuration restricts the maximum number of PCs that can access the Internet by using the same user account. The maximum number of learnable MAC addresses of a service flow is 1023 by default.

----End

Example

The following configurations are used as an example to provision the Internet access service for a user:

- GPON port 0/2/0 on an OLT connects to ONT 1.
- The planned configurations of the user are as follows:
  - The rate of the Internet access service is 4096 Kbit/s.
  - The user connects to Ethernet port 1 on the ONT.
  - The ID of the service VLAN (S-VLAN) is 1000.
  - At most three users can concurrently access the Internet using the same account.
- No proper traffic profile is available and traffic profile **hsi** is created.
- The user has not registered. Do not provision the Internet access service for the user.

```
huawei(config)#traffic table ip name hsi cir 4096 priority 3 priority-policy loc
al-Setting
  Create traffic descriptor record successfully
  -----------------------------------------------
  TD Index           : 10
  TD Name            : hsi
  Priority           : 3
  Copy Priority       : -
  Mapping Index       : -
  CTAG Mapping Priority: -
  CTAG Mapping Index   : -
  CTAG Default Priority: 0
  Priority Policy     : local-pri
```

```
  CIR               : 4096 kbps
  CBS               : 133072 bytes
  PIR               : 8192 kbps
  PBS               : 264144 bytes
  Fix               : 0 kbps
  CAR Threshold Profile: -
  Color Mode         : color-blind
  Color policy       : dei
  Referenced Status   : not used
  ------------------------------------------------
huawei(config)#service-port 5 vlan 1000 port 0/2/0 ont 1 eth 1 multi-service
 user-vlan untagged inbound traffic-table name hsi outbound traffic-table name hsi
huawei(config)#mac-address max-mac-count service-port 5 3
huawei(config)#service-port 5 adminstatus disable
```

The following configurations are used as an example to provision the Internet access service for a commercial user:

- GPON port 0/2/0 on an OLT connects to ONT 2.
- The planned configurations of the user are as follows:
  - The rate of the Internet access service is 8192 Kbit/s.
  - The Internet access service is in multi-service mode to facilitate follow-up service expansion.
  - The user is identified by customer VLAN (C-VLAN).
  - The ID of the S-VLAN is 1023.
  - The ID of the C-VLAN is 100.
  - ONT port 2 connects to the service flow.
- No proper traffic profile is available and traffic profile **huawei** is created.
- The description of the service flow is added to facilitate maintenance.
- The Internet access service needs to be provisioned immediately.

```
huawei(config)#display traffic table ip from-index 0
{ <cr>|to-index<K> }:

  Command:
        display traffic table ip from-index 0
  ------------------------------------------------------------------------------
   TID CIR(kbps) CBS(bytes) PIR(kbps) PBS(bytes) Pri Copy-policy   Pri-Policy
  ------------------------------------------------------------------------------
     0     1024     34768     2048     69536   6 -            tag-pri
     1     2496     81872     4992    163744   6 -            tag-pri
     2      512     18384     1024     36768   0 -           tag-pri
     3      576     20432     1152     40864   2 -           tag-pri
     4       64      4048      128      8096   4 -          tag-pri
     5     2048     67536     4096    135072   0 -            tag-pri
     6      off       off      off       off   0 -          tag-pri
  ------------------------------------------------------------------------------
  Total Num : 7
huawei(config)#traffic table ip name huawei cir 8192 priority 4 priority-policy
 local-Setting
  Create traffic descriptor record successfully
  ------------------------------------------------
  TD Index           : 8
```

```
 TD Name           : huawei
 Priority          : 4
 Copy Priority      : -
 Mapping Index      : -
 CTAG Mapping Priority: -
 CTAG Mapping Index  : -
 CTAG Default Priority: 0
 Priority Policy    : local-pri
 CIR               : 8192 kbps
 CBS               : 264144 bytes
 PIR               : 16384 kbps
 PBS               : 526288 bytes
 Fix               : 0 kbps
 CAR Threshold Profile: -
 Color Mode         : color-blind
 Color policy       : dei
 Referenced Status   : not used
 ------------------------------------------------
huawei(config)#service-port 10 vlan 1023 port 0/2/0 ont 2 eth 2 multi-service
 user-vlan 100 inbound traffic-table name huawei outbound traffic-table name huawei
huawei(config)#service-port desc 10 description gpon/Vlanid:1023/uservlan:100
```

The following configurations are used as an example to provision the Internet access service for a home user:

- Original configuration of the user:
  - The Internet access service is provisioned on ONT port 1.
  - The ID of the service flow is 10.
- New configuration of the user:
  - The Internet access service is provisioned on ONT ports 1-3.

```
huawei(config)#service-port 10 modify ont eth 1-3
```

Configuring Automatic Service Flow Creation (GPON)

The OLT supports configuration of service flow creation policies on PON ports. After an ONU goes online under the PON port, service flows are automatically created according to the preset policies. This topic uses MA5671-24FE as an example for the ONU. Configuration methods are similar for other device models.

Data Plan

Table 13-31 Key data plan

Configuration Item	Data
Networking data	Upstream port: 0/19/0 PON port: 0/1/0 ONU ID: 1
VLAN	Internet access service VLAN (S-VLAN) • S-VLAN ID: 100

Configuration Item	Data
	• S-VLAN type: smart • When the S-VLAN plan is single-tagged VLAN, the S-VLAN attribute is common by default and does not need to be configured. When the S-VLAN plan is double-tagged VLAN, the S-VLAN attribute is QinQ or stacking. A PC is connected to the ONU on the user side. User packets are untagged.
Traffic profile	Profile name: ftto_hsi Committed information rate (CIR): 10 Mbit/s Peak information rate (PIR): 20 Mbit/s Priority policy: user-cos Default 802.1p priority of packets: 0 Enqueue priority policy for scheduling packets: Tag-In-Package
Service flow automatic creation policy	Maximum number of service flows automatically created for each ONU: 24 (When more than 24 service flows are automatically created, service flows can still be created successfully, and at the same time an alarm will be reported indicating that the configuration has exceeded the permitted range.) Upstream traffic profile: ftto_hsi Downstream traffic profile: ftto_hsi VLAN translation policy: • When creating single-tagged service flows, the C-VLAN ranges of all ONUs are the same. The OLT performs VLAN translation by adding an S-VLAN to packets going upstream. This policy is C-VLAN <-> S-VLAN. • When creating double-tagged service flows, the C-VLAN ranges of all ONUs are the same. The OLT performs VLAN translation by adding S-VLAN and translating C-VLAN into C'-VLAN to packets going upstream. The network-side inner VLAN (C'-VLAN) ID starts from 2000. This policy is C-VLAN <-> S-VLAN+C'-VLAN.
DBA profile	Profile name: ftto_dba Profile type: type3 Assured bandwidth: 10 Mbit/s Maximum bandwidth: 20 Mbit/s
ONU line profile	Profile name: ftto_line T-CONT ID: 1 GEM port IDs for Internet access service: 1, 2, 3 **NOTE** Assuming that the service is provisioned to all of the 24 ports of the ONU, each GEM port is mapped to 8 ports. Therefore, 3 GEM ports are needed.
ONU service profile	Profile name: ftto_ser ONU port capability set: 24

Configuration Item	Data
ONU automatic addition policy	Device type: MA5671-F24 Name of ONU line profile: ftto_line Name of ONU service profile: ftto_ser

Procedure

Configure Internet access service VLAN (S-VLAN).

Create Internet access service VLAN 100, and add upstream port 0/19/0 to VLAN 100.

```
huawei(config)#vlan 100 smart
huawei(config)#port vlan 100 0/19 0
```

When a double-tagged VLAN is planned, run the following command; when a single-tagged VLAN is planned, the following command is not required.

```
huawei(config)#vlan attrib 100 q-in-q
```

Step 1 Configure a traffic profile.

You can run the **display traffic table ip** command to query the traffic profiles that exist in the system. If the existing traffic profiles in the system cannot meet the requirements, add a new traffic profile by running the **traffic table ip** command.

```
huawei(config)#traffic table ip name ftto_hsi cir 10240 pir 20480 priority user-cos
0 priority-policy tag-In-package
```

Step 2 Configure the service flow automatic creation policy.

One PON port can be configured with only one policy at a time. The policy can be configured based on PON port or PON board. The following uses the example of configuring a policy based on GPON port 0/1/0.

- Create a single-tagged service flow.

  Create a single-tagged service flow for the ONU. For packets going upstream, a specified S-VLAN is added; for packets going downstream, the VLAN tag is removed.

  ```
  huawei(config)#interface gpon 0/1
  huawei(config-if-gpon-0/1)#auto-service-port port 0 vlan 100 single-vlan
  onu-vlan-num 24 inbound traffic-table name ftto_hsi outbound traffic-table name
  ftto_hsi
  huawei(config-if-gpon-0/1)#quit
  ```

- Create a double-tagged service flow.

  Create a service flow for each port of the ONU. For packets going upstream, C'-VLAN and S-VLAN are added. The C'-VLAN needs to be calculated based on the start VLAN ID, ONU ID, and port ID. A C'-VLAN is calculated using the following formula: Start VLAN ID + ONU ID x ONU port quantity + ONU user port ID. For example, if the start VLAN ID is 2000, ONU ID is 1, ONU port quantity is 24, and ONU user port ID is 1, the C'-VLAN ID is 2025 (2000 + 1 x 24 + 1). For the 24 Ethernet ports of the ONU, the value range of C'-VLAN is 2025-2048.

```
huawei(config)#interface gpon 0/1
huawei(config-if-gpon-0/1)#auto-service-port port 0 vlan 100 double-vlan from-vlan 2000 onu-vlan-num 24 inbound traffic-table name ftto_hsi outbound traffic-table name ftto_hsi
huawei(config-if-gpon-0/1)#quit
```

Step 3 Configure GPON ONU profiles.

GPON ONU profiles include the DBA profile, line profile, service profile, and alarm profile.

- DBA profile: A DBA profile describes GPON traffic parameters. A T-CONT is bound to a DBA profile for dynamic bandwidth allocation, improving upstream bandwidth utilization.
- Line profile: A line profile describes the binding between a T-CONT and a DBA profile, the QoS mode of service flows, and the mapping between GEM ports and the ONU-side service.
- Service profile: A service profile provides the service configuration channel for the ONU that is managed through OMCI.
- Alarm profile: An alarm profile contains a series of alarm thresholds to measure and monitor the performance of activated ONU lines. When a statistical value reaches the threshold, the host is notified and reports an alarm to the log host and the NMS.

1. Configure a DBA profile.

   You can run the **display dba-profile** command to query the existing DBA profiles in the system. If the existing DBA profiles in the system cannot meet the requirements, add a new DBA profile by running the **dba-profile add** command.

   ```
   huawei(config)#dba-profile add profile-name ftto_dba type3 assure 10240 max 20480
   ```

   NOTE
   The DBA implementation is based on an ONU. Therefore, select a DBA profile of the proper bandwidth type and configure proper bandwidths according to the service types and total user count of the ONU. Note that the sum of the fixed bandwidth and the assured bandwidth must not be greater than the total bandwidth of the PON port.

2. Configure an ONU line profile.

   Create a GPON ONU line profile named **ftto_line**, and bind the line profile to DBA profile **ftto_dba**.

   ```
   huawei(config)#ont-lineprofile gpon profile-name ftto_line
   huawei(config-gpon-lineprofile-1)#tcont 1 dba-profile-name ftto_dba
   ```

   Bind GEM ports 1, 2, and 3 to T-CONT 1.

   ```
   huawei(config-gpon-lineprofile-1)#gem add 1 eth tcont 1
   huawei(config-gpon-lineprofile-1)#gem add 2 eth tcont 1
   huawei(config-gpon-lineprofile-1)#gem add 3 eth tcont 1
   ```

   Configure the mapping between the 3 GEM ports and the 24 C-VLANs.

   - When user packets are untagged and MA5671 is automatically added, each Ethernet port of MA5671 will be marked with different native VLANs. The native VLAN ID equals the Ethernet port ID.
   - When user packets carry a C-VLAN tag, configure the corresponding VLAN mapping in the line profile, and configure the corresponding VLAN in the service profile. That is, in the following configuration, replace the VLAN ID with the corresponding C-VLAN ID. For example, when the C-VLAN ID of the user packets received by Ethernet port 1 of MA5671 is 1001, the command should be **gem mapping 1 0 vlan 1001**.

```
huawei(config-gpon-lineprofile-1)#gem mapping 1 0 vlan 1
huawei(config-gpon-lineprofile-1)#gem mapping 1 1 vlan 2
huawei(config-gpon-lineprofile-1)#gem mapping 1 2 vlan 3
huawei(config-gpon-lineprofile-1)#gem mapping 1 3 vlan 4
huawei(config-gpon-lineprofile-1)#gem mapping 1 4 vlan 5
huawei(config-gpon-lineprofile-1)#gem mapping 1 5 vlan 6
huawei(config-gpon-lineprofile-1)#gem mapping 1 6 vlan 7
huawei(config-gpon-lineprofile-1)#gem mapping 1 7 vlan 8
huawei(config-gpon-lineprofile-1)#gem mapping 2 0 vlan 9
huawei(config-gpon-lineprofile-1)#gem mapping 2 1 vlan 10
huawei(config-gpon-lineprofile-1)#gem mapping 2 2 vlan 11
huawei(config-gpon-lineprofile-1)#gem mapping 2 3 vlan 12
huawei(config-gpon-lineprofile-1)#gem mapping 2 4 vlan 13
huawei(config-gpon-lineprofile-1)#gem mapping 2 5 vlan 14
huawei(config-gpon-lineprofile-1)#gem mapping 2 6 vlan 15
huawei(config-gpon-lineprofile-1)#gem mapping 2 7 vlan 16
huawei(config-gpon-lineprofile-1)#gem mapping 3 0 vlan 17
huawei(config-gpon-lineprofile-1)#gem mapping 3 1 vlan 18
huawei(config-gpon-lineprofile-1)#gem mapping 3 2 vlan 19
huawei(config-gpon-lineprofile-1)#gem mapping 3 3 vlan 20
huawei(config-gpon-lineprofile-1)#gem mapping 3 4 vlan 21
huawei(config-gpon-lineprofile-1)#gem mapping 3 5 vlan 22
huawei(config-gpon-lineprofile-1)#gem mapping 3 6 vlan 23
huawei(config-gpon-lineprofile-1)#gem mapping 3 7 vlan 24
```

After completing the configuration, run the **commit** command to make the configured parameters take effect.

```
huawei(config-gpon-lineprofile-1)#commit
huawei(config-gpon-lineprofile-1)#quit
```

3. Configure an ONU service profile.

   Create a GPON ONU service profile named **ftto_ser**. Set the Ethernet port capability set to 24.

```
huawei(config)#ont-srvprofile gpon profile-name ftto_ser
huawei(config-gpon-srvprofile-1)#ont-port eth 24
```

   When the user packets carry a C-VLAN tag, you need to replace the VLAN ID in the following configuration with the corresponding C-VLAN ID. For example, when the C-VLAN ID of the user packets received by Ethernet port 1 of MA5671 is 1001, the command should be **port vlan eth 1 1001**.

```
huawei(config-gpon-srvprofile-1)#port vlan eth 1 1
huawei(config-gpon-srvprofile-1)#port vlan eth 2 2
huawei(config-gpon-srvprofile-1)#port vlan eth 3 3
huawei(config-gpon-srvprofile-1)#port vlan eth 4 4
huawei(config-gpon-srvprofile-1)#port vlan eth 5 5
huawei(config-gpon-srvprofile-1)#port vlan eth 6 6
huawei(config-gpon-srvprofile-1)#port vlan eth 7 7
huawei(config-gpon-srvprofile-1)#port vlan eth 8 8
huawei(config-gpon-srvprofile-1)#port vlan eth 9 9
huawei(config-gpon-srvprofile-1)#port vlan eth 10 10
huawei(config-gpon-srvprofile-1)#port vlan eth 11 11
huawei(config-gpon-srvprofile-1)#port vlan eth 12 12
huawei(config-gpon-srvprofile-1)#port vlan eth 13 13
huawei(config-gpon-srvprofile-1)#port vlan eth 14 14
huawei(config-gpon-srvprofile-1)#port vlan eth 15 15
huawei(config-gpon-srvprofile-1)#port vlan eth 16 16
```

```
huawei(config-gpon-srvprofile-1)#port vlan eth 17 17
huawei(config-gpon-srvprofile-1)#port vlan eth 18 18
huawei(config-gpon-srvprofile-1)#port vlan eth 19 19
huawei(config-gpon-srvprofile-1)#port vlan eth 20 20
huawei(config-gpon-srvprofile-1)#port vlan eth 21 21
huawei(config-gpon-srvprofile-1)#port vlan eth 22 22
huawei(config-gpon-srvprofile-1)#port vlan eth 23 23
huawei(config-gpon-srvprofile-1)#port vlan eth 24 24
```

After completing the configuration, run the **commit** command to make the configured parameters take effect.

```
huawei(config-gpon-srvprofile-1)#commit
huawei(config-gpon-srvprofile-1)#quit
```

4. (Optional) Configure an alarm profile.
   - The ID of the default GPON alarm profile in the system is 1. The thresholds of all the alarm parameters in the default alarm profile are 0, which indicates that no alarm is reported.
   - In this example, the default alarm profile is used, and therefore the configuration of the alarm profile is not required.
   - To configure the alarm threshold parameters for monitoring the performance statistics of an activated ONU line, run the **gpon alarm-profile add** command to configure a GPON alarm profile.

Step 4 Configure an ONU automatic addition policy.

The **equipment** parameter needs to be determined by the actual equipment type. Here, the 24-port MA5671 is used as an example. The value of **equipment** is **MA5671-F24**. For a 16-port MA5671, the value of **equipment** is **MA5671-F16**; for an 8-port MA5671, the value of **equipment** is **MA5671-F8**. The same policy cannot be applied to different types of ONUs. Each type of ONU requires a corresponding policy.

When **auto-confirm** is set to **enable**, you do not need to run the **ont confirm** command on the GPON port to confirm the ONU. When **auto-confirm** is set to **disable**, you need to run the **ont confirm** command on the GPON port.

```
huawei(config)#ont auto-add-policy gpon equipment MA5671-F24 omci ont-lineprofile-name
ftto_line ont-srvprofile-name ftto_ser auto-confirm enable
```

After completing the configuration, run the **display ont auto-add-policy** command to query the ONU automatic addition policy.

```
huawei(config)#display ont auto-add-policy gpon all
------------------------------------------------------------------------------
  Index               : 1
  Ont EquipmentID      : MA5671-F24(0x4d41353637312d463234)
  PON mode            : GPON
  Management mode      : OMCI
  Line profile ID      : 1
  Line profile name     : ftto_line
  Service profile ID    : 1
  Service profile name   : ftto_ser
  Auto-confirm         : enable
  ------------------------------------------------------------------------------
```

Step 5 Add the ONU.

1. Enable the ONU auto-find function on the GPON port.

   After you have configured the ONU automatic addition policy, you do not need to add the ONU offline. The ONU will be automatically discovered after going online.

   ```
   huawei(config)#interface gpon 0/1
   huawei(config-if-gpon-0/1)#port 0 ont-auto-find enable
   ```

2. (Optional) Confirm the ONU.

   When **auto-confirm** in the ONU automatic addition policy is set to **disable**, you need to run the **ont-confirm** command to confirm the ONU. If you do not specify the line profile and service profile when running the **ont confirm** command, and the ONU type matches the ONU type configured in the automatic addition policy, the ONU will be automatically bound to the line profile and service profile configured in the automatic addition policy. If the ONU types do not match, the ONU will be bound to default line profile 0.

   ```
   huawei(config-if-gpon-0/1)#display ont autofind 0    /*This command displays information about all the ONTs connected to the GPON port through optical splitters.*/
   huawei(config-if-gpon-0/1)#ont confirm 0 ontid 1 sn-auth 3230313126595540 omci
   ```

3. (Optional) Modify the ONU description.

   You can run the **ont modify** *portid ontid* **desc** *describe-value* command to modify the ONU description. It is advised to add description such as the location and time for each ONU to facilitate fault locating and maintenance.

   ```
   huawei(config-if-gpon-0/1)#ont modify 0 1 desc ftto_2015
   ```

4. (Optional) Bind the alarm profile to the ONU.

   The default alarm profile (profile 1) is used in this example.

   ```
   huawei(config-if-gpon-0/1)#ont alarm-profile 0 1 profile-id 1
   ```

5. Check the ONU status.

   After adding an ONU, run the **display ont info** command to query the current ONU status. Ensure that values for **Control flag**, **Run State**, **Config state**, and **Match state** of the ONU are **active**, **online**, **normal**, and **match** respectively.

   ```
   huawei(config-if-gpon-0/1)#display ont info 0 1
     ------------------------------------------------------------------------
     F/S/P             : 0/1/0
     ONT-ID            : 1
     Control flag       : active   //Indicates that the ONU is activated.
     Run state         : online   //Indicates that the ONU has gone online.
     Config state       : normal   //Indicates that the ONU configuration recovery
   is normal.
     Match state        : match    //Indicates that the capability profile bound to
   the ONU is consistent with the capabilities supported by the ONU.
   ...//The subsequent display is omitted.
   ```

   When ONU configuration fails or the ONU fails to enter the up state:

   - If **Control flag** is **deactive**, run the **ont activate** command in GPON port mode to activate the ONU.
   - If the ONU fails to enter the up state, that is, **Run state** is **offline**, a physical line may be broken or the optical module may be damaged. Check the line and the optical module.
   - If ONU configuration fails, that is, **Config state** is **failed**, the configured ONU capability set exceeds the capabilities supported by the ONU. In such a case, run the

display ont failed-configuration command in diagnose mode to check the failed configuration item and the failure causes. Rectify the faults accordingly.

Step 6 Query the service flow configuration status.

After MA5671-FE24 powers up and connects to the PON network, service flows will be automatically created in about 1 minute. On the OLT, run the **display service-port all** command to query whether the service flows are successfully created.

```
huawei(config)#display service-port all
 Switch-Oriented Flow List
 -----------------------------------------------------------------------------
  INDEX VLAN VLAN    PORT F/ S/ P VPI  VCI   FLOW  FLOW      RX   TX   STATE
        ID   ATTR    TYPE                TYPE  PARA
 -----------------------------------------------------------------------------
     0  100 common   gpon 0/1 /0  1    1    vlan  1         7    7    up
     1  100 common   gpon 0/1 /0  1    1    vlan  2         7    7    up
     ...
     8  100 common   gpon 0/1 /0  1    2    vlan  9         7    7    up
     9  100 common   gpon 0/1 /0  1    2    vlan  10        7    7    up
     ...
    16  100 common   gpon 0/1 /0  1    3    vlan  17        7    7    up
    17  100 common   gpon 0/1 /0  1    3    vlan  18        7    7    up
     ...
```

Meanings of the main parameters are as follows:

- **VPI** indicates the ONU ID.
- **VCI** indicates the GEM port ID of the GPON port.
- **FLOW PARA** indicates the user VLAN ID.

When creating a double-tagged service flow, you can run the **display service-port** *index* command to query the network-side inner VLAN. For example, the query results are as follows when a service flow is automatically created with start VLAN ID 2000, ONU ID 1, ONU port quantity 24, and ONU user port ID 1. Among the parameters, **Label** indicates the network-side inner VLAN.

```
huawei(config)#display service-port 0
 ------------------------------------------------------------
 Index             : 0
 VLAN ID           : 100
 VLAN attr         : QinQ
 Port type         : gpon
 F/S/P             : 0/1/0
 ONT ID            : 1
 GEM port index     : 1
 ONT Port          : 1
 Flow type         : vlan
 Flow para         : 1
 TX               : 7
 Inbound table name  : ftto_hsi
 RX               : 7
 Outbound table name : ftto_hsi
 Admin status       : enable
 State             : up
 Label             : 2025
 Priority          : 0
```

```
PVC bundle        : no
Max MAC count     : 1023
Tag transform     : translate-and-add
Description       :
Remote description  :
Service-port bundle : -
Cos              : -
Car-Group         : -
Static MAC        :
IP address        :
-------------------------------------------------------------
```

----End

13.4.4 Maintenance and Diagnosis

Query the MAC address of a user and the packet statistics of a service flow on an access node to determine whether the service flow is functional.

A service flow is a Layer 2 forwarding channel between an access node and a user. Check the status of a service flow using the following methods:

- After a user goes online, run the **display mac-address** command to query the MAC address of the user. This method is used to check whether the configuration of the service flow is correct.
- After a service is provisioned, run the **display statistics service-port** command to query the packet statistics of the service flow or run the **display traffic service-port** command to query the packet transmit and receive rate of the service flow in real time. The query results can be used to detect a service exception or locate a fault.

13.4.5 Reference Standards and Protocols

None

13.5 Service Port Bundle

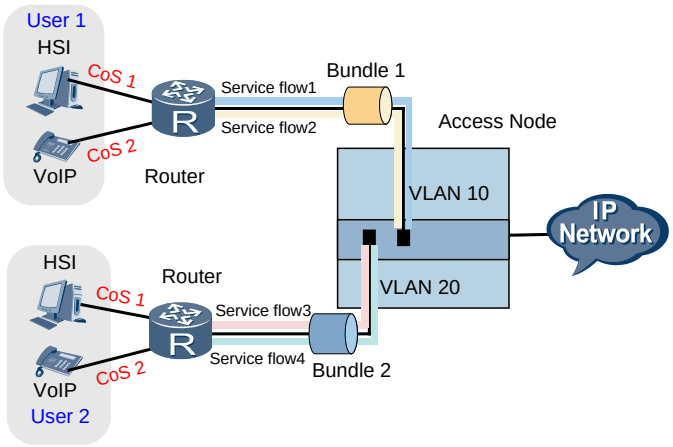
13.5.1 What Is Service Port Bundle

Definition

In a network where VLANs are planned based on ports on the access node and the access node connects to users through a router, the packets transmitted from a user to the access node through the router may carry the same MAC address even when they carry different services. As a result, MAC address transfer may occur between different service flows (service flow 1 and service flow 2) on the access node. The service port bundle feature addresses this issue by using CoS-based route selection.

Service port bundle, also named service flow bundle or flow bundle, is a CoS-based packet forwarding model. Each service port bundle corresponds to a group of services for a user. Each service flow carries one type of service and has a CoS level. A service port bundle can also be considered a bundle of service flows. The following figure shows the schematic diagram for service port bundle.

Figure 13-21 Schematic diagram for service port bundle



Benefits

Benefits to carriers

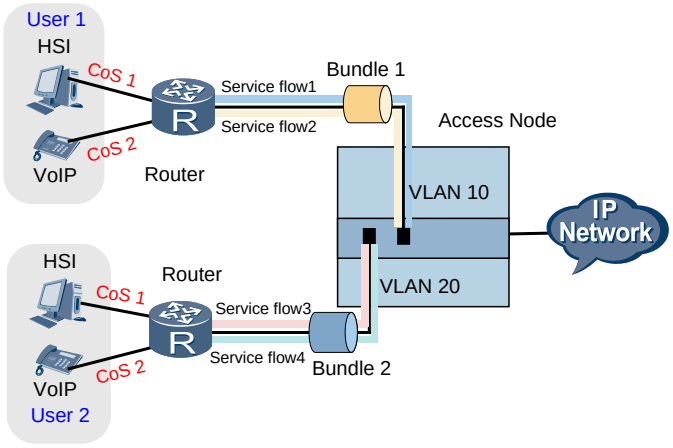
CoS identification is introduced to the service port bundle principle. Consequently, service port bundle breaks the VLAN+MAC and S-VLAN+C-VLAN forwarding policies for the access node down to the VLAN+MAC+CoS and S-VLAN+C-VLAN+CoS forwarding policies. and improves the flexibility in service planning for carriers.

13.5.2 Schematic Diagram For Service Port Bundle

The following explains the schematic diagram for service port bundle.

- The access node connects to users through a router. CoS levels are used to distinguish different services of a user.
- VLANs on the access node are planned based on ports. Services of different users are carried by different service flows. Service flows of a user are bundled together, called the service port bundle, or service flow bundle.

Figure 13-22 Schematic diagram for service port bundle



Upstream direction

The Access node receives packets of the service port bundle, and forwards the received packets to the upper-layer network through upstream ports in the S-VLAN (in the above figure, service flows 1 and 2 in service port bundle 1 are converged to the upstream port in VLAN 10). At the same time, the Access node obtains the service port bundle identifier from the ARL table (learned from the user port) and creates an entry for the service port bundle.

Downstream direction

The Access node determines the user port based on S-VLAN+MAC or S-VLAN+C-VLAN, queries the service port bundle table, determines a service flow based on the CoS value, and forwards the received packets through the service flow. In this way, the Access node implements the VLAN+MAC+CoS or S-VLAN+C-VLAN+CoS forwarding policy.

For more information about Layer 2 forwarding policies, see 13.6 Layer 2 Forwarding Policy.

13.5.3 Configuring a Service Port Bundle

On an access node, VLANs are planned based on ports. If the access node connects to users through a router, configure a service port bundle to prevent MAC address flapping between various service flows of the users.

Procedure

Run the **service-port-bundle** command to create a service port bundle.

A service port bundle corresponds to a group of services. It is recommended that you plan a unique service port bundle ID for a user.

Step 1 Add service flows to the service port bundle and set classes of service (CoSs) for the service flows.

The services in a service flow bundle are carried by service flows. Add the service flows to the service port bundle and set a CoS for each service flow. In this way, the router forwards service packets according to the CoS.

Run the **service-port(profile-mode)** or **rvice-port(distributing-mode)** command to set **bundle** *bundleid* **cos** *cos* when creating a service flow. *bundleid* is the ID of the service port bundle created in the preceding step.

----End

Example

The following is an example of the configurations used to configure a service port bundle to prevent MAC address flapping between Internet access and multicast service flows of user A:

- Ethernet port 0/2/1 on the access node connects to user A through a router.
- The VLANs of the access node are planned based on ports.
- Ethernet port 0/2/1 is in VLAN 10.
- User A requires the Internet access and broadcast services.
- The ID of the service port bundle for user A is 10.
- The Internet access and broadcast services of user A are differentiated based on CoS values.
- For the Internet access service:

  - The CoS value is 1.
  - The ID of the traffic profile is 10.
- For the broadcast service:
  - The CoS value is 3.
  - The ID of the traffic profile is 20.

```
huawei(config)#service-port-bundle 10
huawei(config)#service-port 10 vlan 10 eth 0/2/1 bundle 10 cos 1 inbound
traffic-table index 10 outbound traffic-table index 10
huawei(config)#service-port 11 vlan 10 eth 0/2/1 bundle 10 cos 3 inbound
traffic-table index 20 outbound traffic-table index 20
```

13.6 Layer 2 Forwarding Policy

13.6.1 Overview

The MA5600T/MA5603T/MA5608T, as Layer 2 network equipment, supports the ability to transparently transmit or forward packets at Layer 2. Traditionally, Layer 2 packet forwarding is based on VLANs and MAC addresses of packets (VLAN+MAC). If the destination MAC address of a packet becomes ineffective due to dynamic MAC address aging, VLAN+MAC searching fails. The packet becomes an unknown unicast packet and is broadcast within the VLAN, which poses a security threat. In addition, VLAN+MAC forwarding is subject to MAC spoofing and attacks, which lead to security problems.

To address the preceding problems, you can use S-VLAN+C-VLAN forwarding instead.

In S-VLAN+C-VLAN forwarding, two VLAN tags form a Layer 2 forwarding mapping relationship. Packets are forwarded based on VLANs rather than MAC addresses.

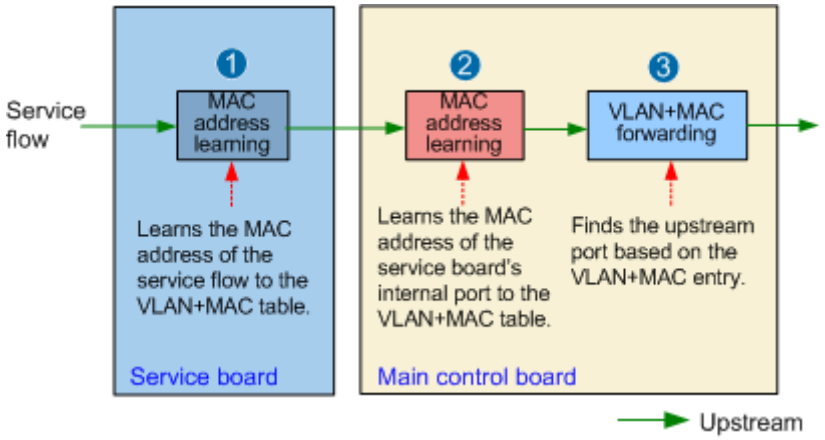
13.6.2 Principles

The MA5600T/MA5603T/MA5608T supports VLAN+MAC forwarding, S-VLAN+C-VLAN forwarding, and two other forwarding modes derived from the first two: VLAN+MAC+CoS and S-VLAN+C-VLAN+CoS.

VLAN+MAC Forwarding

Figure 13-23 and Figure 13-24 show a VLAN+MAC forwarding process.

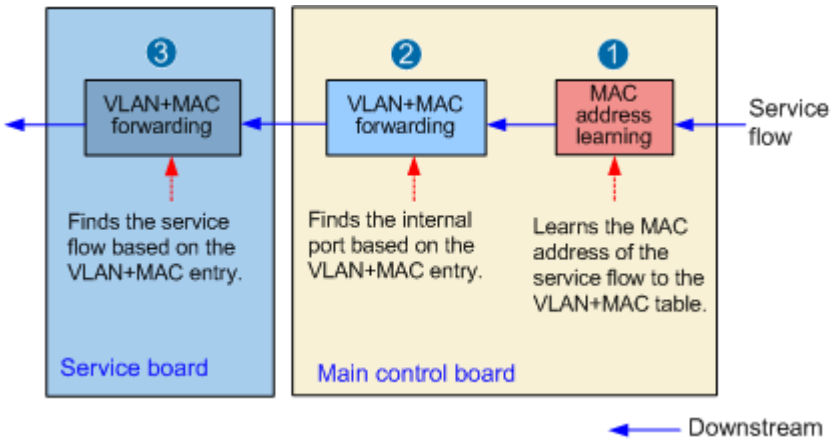
Figure 13-23 VLAN+MAC forwarding model in the upstream direction



Upstream direction

1. The service board learns the MAC address of a service flow to the VLAN+MAC table so that the service flow can be found based on the VLAN+MAC entry during downstream forwarding.
2. The main control board learns the MAC address of the service board's internal port to the VLAN+MAC table so that the internal port of the service board can be found based on the VLAN+MAC entry during downstream forwarding.
3. The main control board finds the upstream port and forwards the packets based on the VLAN+MAC forwarding entry.

Figure 13-24 VLAN+MAC forwarding model in the downstream direction



Downstream direction

1. The main control board learns the MAC address of an upstream port to the VLAN+MAC table so that the upstream port can be found based on the VLAN+MAC entry during upstream forwarding.
2. The main control board finds the internal port of the service board based on the VLAN+MAC forwarding entry.

3. The service board finds the service flow and forwards the packets based on the VLAN+MAC forwarding entry.

NOTE

- In the upstream direction, the service board does not forward packets based on VLAN+MAC, but forwards packets directly to the main control board. Therefore, in the downstream direction, the service board does not need to learn the MAC address of the main control board.
- In the VLAN+MAC forwarding mechanism, if a packet carries a broadcast MAC address or unknown unicast MAC address, the packet is broadcast in the VLAN. That is, the packet is duplicated and transmitted to every port in the VLAN.

S-VLAN+C-VLAN Forwarding

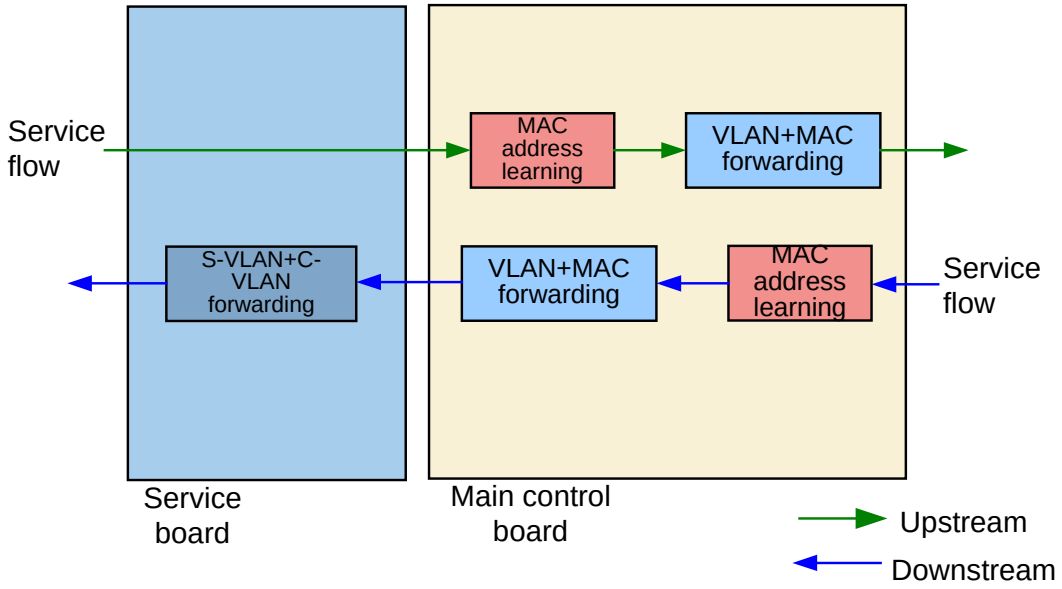
The two VLANs (S-VLAN and C-VLAN) are an extension of VLAN. This expands the VLAN ID range. In addition, S and C has different meanings. S stands for service, and C stands for customer (user). Each S-VLAN+C-VLAN uniquely identifies a user service, and S-VLAN+C-VLAN forwarding can be implemented.

In S-VLAN+C-VLAN forwarding, two VLAN tags form a Layer 2 forwarding mapping relationship to implement VLAN-based forwarding. S-VLAN+C-VLAN forwarding entries do not need to be learned dynamically. The system automatically creates static forwarding entries during the establishment of service flows. According to the forwarding entries, upstream packets are transmitted through the corresponding upstream port and downstream packets are transmitted through the corresponding service port. S-VLAN+C-VLAN forwarding is classified into non-strict S-VLAN+C-VLAN forwarding and strict S-VLAN+C-VLAN forwarding.

Non-strict S-VLAN+C-VLAN forwarding

The main control board forwards packets based on VLAN+MAC, and the service board forwards packets based on S-VLAN+C-VLAN, as shown in Figure 13-25.

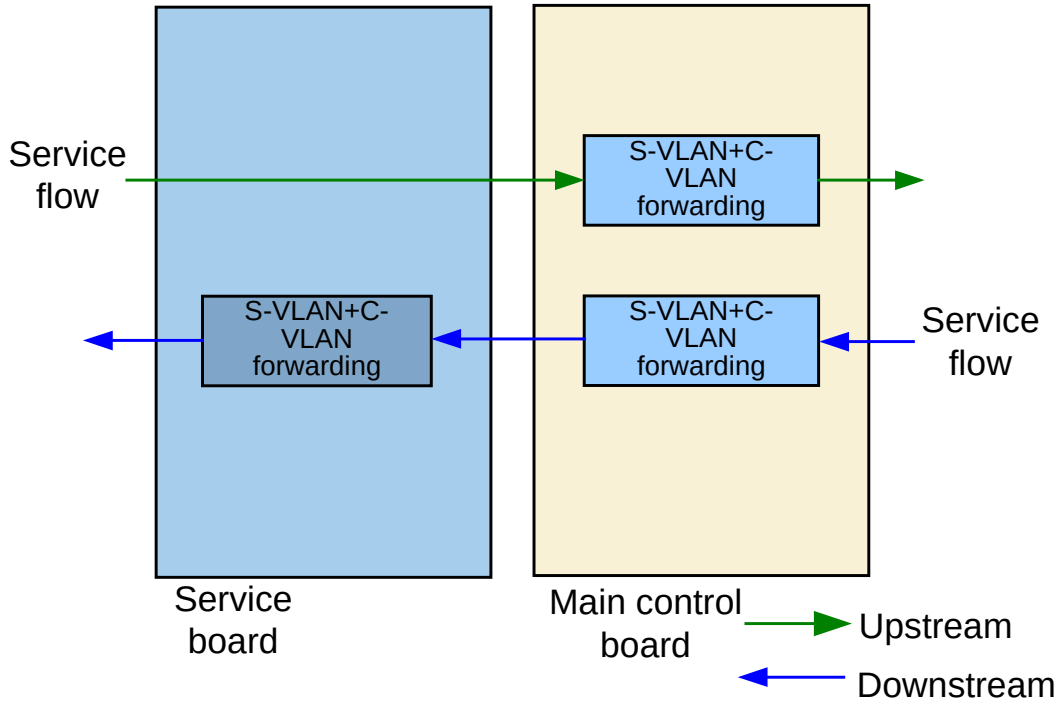
Figure 13-25 Non-strict S-VLAN+C-VLAN forwarding model



Strict S-VLAN+C-VLAN forwarding

The main control and service boards forward packets based on S-VLAN+C-VLAN, and the main control board does not learn MAC addresses, as shown in Figure 13-26.

Figure 13-26 Strict S-VLAN+C-VLAN forwarding model



NOTE

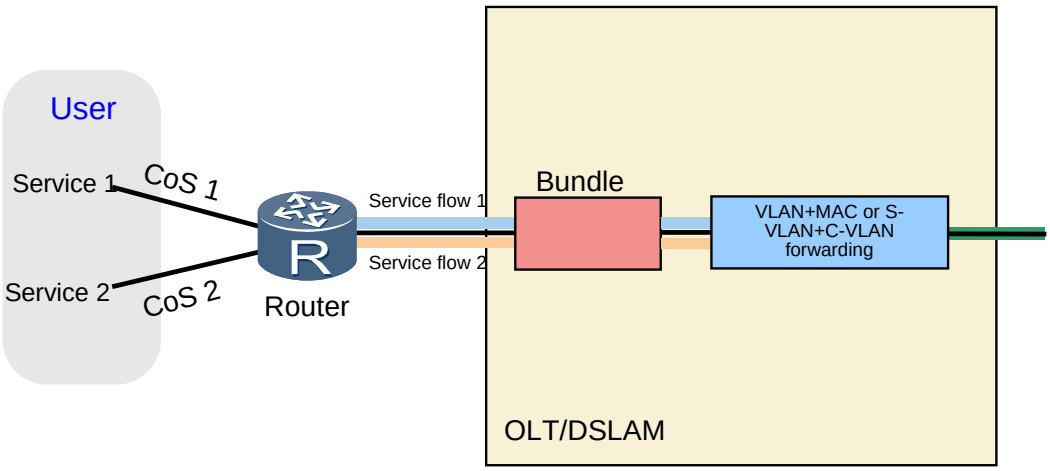
- In the upstream direction, the service board does not forward packets based on S-VLAN+C-VLAN, but forwards packets directly to the main control board.
- In non-strict S-VLAN+C-VLAN forwarding, if the number of MAC addresses learned by the main control board reaches the maximum number, a serious hash conflict occurs, or MAC addresses are aged, a large number of unknown unicast or multicast packets may become broadcast packets, which will occupy many forwarding bandwidth resources. Therefore, on a network with a large number of users (services without a limit on the number of learned MAC addresses), strict S-VLAN+C-VLAN forwarding is recommended.

VLAN+MAC+CoS and S-VLAN+C-VLAN+CoS Forwarding

VLAN+MAC+CoS and S-VLAN+C-VLAN+CoS forwarding policies are derived from VLAN+MAC and S-VLAN+C-VLAN forwarding policies of the MA5600T/MA5603T/MA5608T respectively after class of service (CoS) identification is introduced based on service port bundles. The forwarding policies based on CoS resolve the problem of different service flows with the same VLAN+MAC or S-VLAN+C-VLAN and therefore improve service planning flexibility.

Figure 13-27 shows the CoS-based forwarding model.

Figure 13-27 VLAN+MAC+CoS and S-VLAN+C-VLAN+CoS forwarding model



Upstream direction

All service flows from the same user converge to a service port bundle. The MA5600T/MA5603T/MA5608T learns the service port bundle identifier of the service flows, creates a service port bundle entry, and performs Layer 2 packet forwarding based on VLAN+MAC or S-VLAN+C-VLAN.

Downstream direction

The MA5600T/MA5603T/MA5608T determines a unique user based on VLAN+MAC or S-VLAN+C-VLAN, queries the service port bundle entries, finds the service flows based on CoS, and forwards packets.

For details about service port bundles, see 13.5 Service Port Bundle.

13.6.3 Configuring a Layer 2 Forwarding Policy

Layer 2 forwarding policy configuration is the basis of VLAN configuration and an important step to ensure correct service forwarding. Before configuring services, ensure that the VLAN forwarding policy has been configured as planned.

Prerequisites

- VLAN IDs have been created correctly.
- VLAN attributes have been set correctly.

Context

A VLAN forwarding policy is a type of Layer 2 forwarding policy. Two types of VLAN forwarding policy are available: **vlan-connect** or **vlan-mac**. **vlan-mac** indicates VLAN+MAC forwarding and **vlan-connect** indicates S-VLAN+C-VLAN forwarding.

To implement VLAN+MAC+CoS or S-VLAN+C-VLAN+CoS forwarding, configure the VLAN forwarding policy and service port bundles separately. For details about how to configure service port bundles, see 13.5.3 Configuring a Service Port Bundle.

Default Configuration

Table 13-32 lists the default VLAN forwarding policy settings.

Table 13-32 Default VLAN forwarding policy settings

Parameter	Default Value
forwarding-mode	**vlan-mac**

Procedure

- Configure VLAN+MAC forwarding and non-strict S-VLAN+C-VLAN forwarding.

  You can configure a VLAN forwarding policy using two methods. Select an appropriate one as follows:

  - To configure a VLAN forwarding policy for a single VLAN, use the first method.
  - To configure the same forwarding policy for multiple VLANs with the same service profile parameters, use the second method (batch configuration using a service profile).
  - Method 1

  a. In global configuration mode, run the **vlan forwarding** command. If *forwarding-mode* is **vlan-mac**, VLAN+MAC forwarding is used. If *forwarding-mode* is **vlan-connect**, non-strict S-VLAN+C-VLAN forwarding is used.

  - Method 2

  a. Run the **vlan service-profile** command to create a VLAN service profile and enter the VLAN service profile mode.

  b. Run the **forwarding** command to configure a VLAN forwarding policy. If the VLAN forwarding policy is **vlan-mac**, VLAN+MAC forwarding is used. If the VLAN forwarding policy is **vlan-connect**, non-strict S-VLAN+C-VLAN forwarding is used.

  c. Run the **commit** command for the profile configuration to take effect.

  d. Run the **quit** command to quit the VLAN service profile mode.

  e. Run the **vlan bind service-profile** command to bind a VLAN to the VLAN service profile.

- Configure strict S-VLAN+C-VLAN forwarding.

  - Method 1

  a. Run the **vlan** command to add a VLAN.

  b. Run the **vlan attrib** command to change the VLAN attribute to QinQ or stacking.

  c. Run the **mac-address learning** command to disable MAC address learning on the main control board.

  d. Run the **vlan forwarding** or **forwarding** command to configure a VLAN forwarding policy.

  e. Run the **port vlan** command to associate upstream ports with S-VLANs.

  f. Run the **service-port** *[index]* [**uplink-port** *frameid/slotid/portid*] command to create a VLAN connection between a service port and upstream port.

  - Method 2

a. Run the **vlan** command to add a VLAN.
b. Run the **vlan attrib** command to change the VLAN attribute to QinQ or stacking.
c. Run the **mac-address learning** command to disable MAC address learning on the main control board.
d. Run the **vlan forwarding** or **forwarding** command to configure a VLAN forwarding policy.
e. Run the **port vlan** command to associate upstream ports with S-VLANs and C-VLANs.
f. Run the **service-port** command to create a service port.

----End

Result

After the configuration, service flows can be created based on the configured VLANs and users connected to the MA5600T/MA5603T/MA5608T can ping the upstream equipment.

Example

Example: Configure the forwarding policy of VLAN 50 as S-VLAN+C-VLAN.

```
huawei(config)#vlan forwarding 50 vlan-connect
```

Example: Use a VLAN service profile to configure the forwarding policy of VLAN 60 as S-VLAN+C-VLAN.

```
huawei(config)#vlan service-profile profile-id 10
huawei(config-vlan-srvprof-10)#forwarding vlan-connect
huawei(config-vlan-srvprof-10)#commit
huawei(config-vlan-srvprof-10)#quit
huawei(config)#vlan bind service-profile 60 profile-id 10
```

Example: When the SCUN board is used, configure the forwarding policy of VLAN 65 as strict S-VLAN+C-VLAN. VLAN 73 indicates an enterprise VLAN on which MAC address learning is disabled.

```
huawei(config)#vlan 65 smart
huawei(config)#vlan attrib 65 stacking
huawei(config)#vlan service-profile profile-id 200
huawei(config-vlan-srvprof-200)#mac-address learning fabric disable
huawei(config-vlan-srvprof-200)#forwarding vlan-connect
huawei(config-vlan-srvprof-200)#commit
huawei(config-vlan-srvprof-200)#quit
huawei(config)#vlan bind service-profile 65 profile-id 200
huawei(config)#port vlan 65 inner-vlan-list 73 0/19 0
huawei(config)#service-port 100 uplink-port 0/19/0 vlan 65 gpon 0/2/0 ont 1 gemport
2 multi-service user-vlan 73 rx-cttr 10 tx-cttr 10
```

13.6.4 Reference Standards and Protocols

The reference protocol of the Layer 2 forwarding policy feature is as follows:

DSL Forum TR-101: Migration to Ethernet-Based DSL Aggregation

13.7 Layer 2 User Bridging

Layer 2 user bridging implements Layer 2 data exchange among users under the same MA5600T/MA5603T/MA5608T.

13.7.1 Overview

Users under the same MA5600T/MA5603T/MA5608T are isolated at Layer 2, but QinQ service deployment requires users communicate with each other at Layer 2. Therefore, QinQ services can be deployed only between different MA5600T/MA5603T/MA5608Ts.

If users in the same IP network segment on the same MA5600T/MA5603T/MA5608T need to exchange data, upstream equipment should support ARP proxy to implement Layer 3 data forwarding.

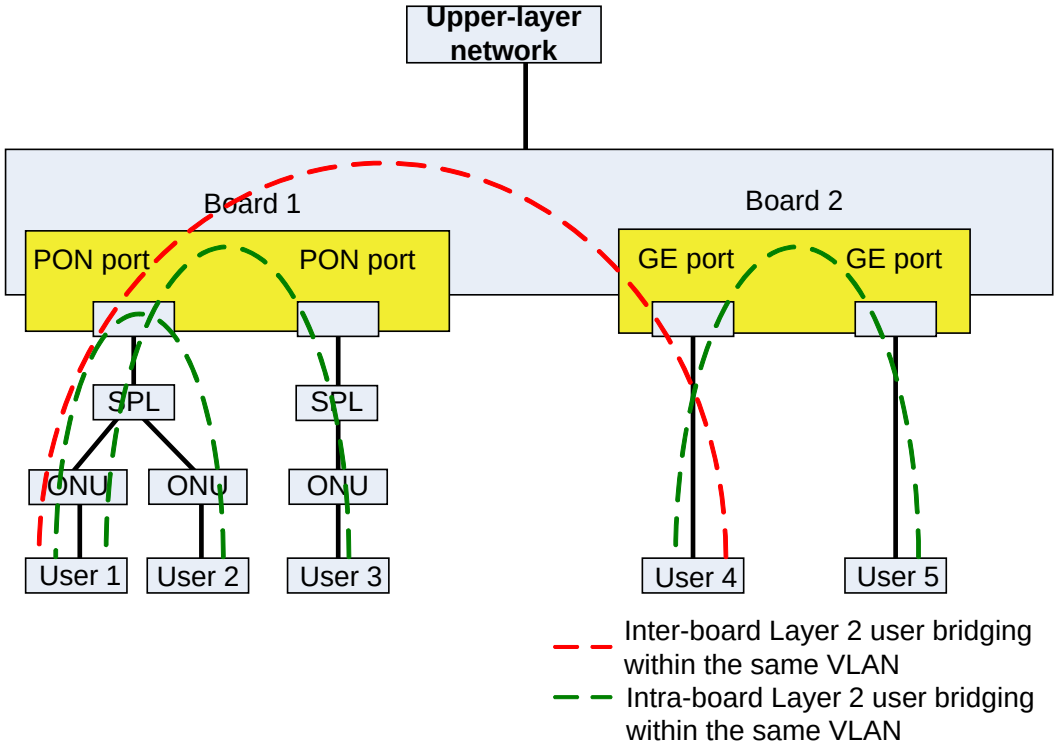
The Layer 2 user bridging feature addresses the preceding issues. After Layer 2 user bridging is enabled, users under the same MA5600T/MA5603T/MA5608T can exchange data at Layer 2.

13.7.2 Principles

Layer 2 user bridging includes VLAN-based and global Layer 2 user bridging mode. Each mode can be further classified into intra-board and inter-board user bridging. For details about the support of different boards for VLAN-based and global Layer 2 user bridging, see Specifications.

VLAN-based Layer 2 User Bridging

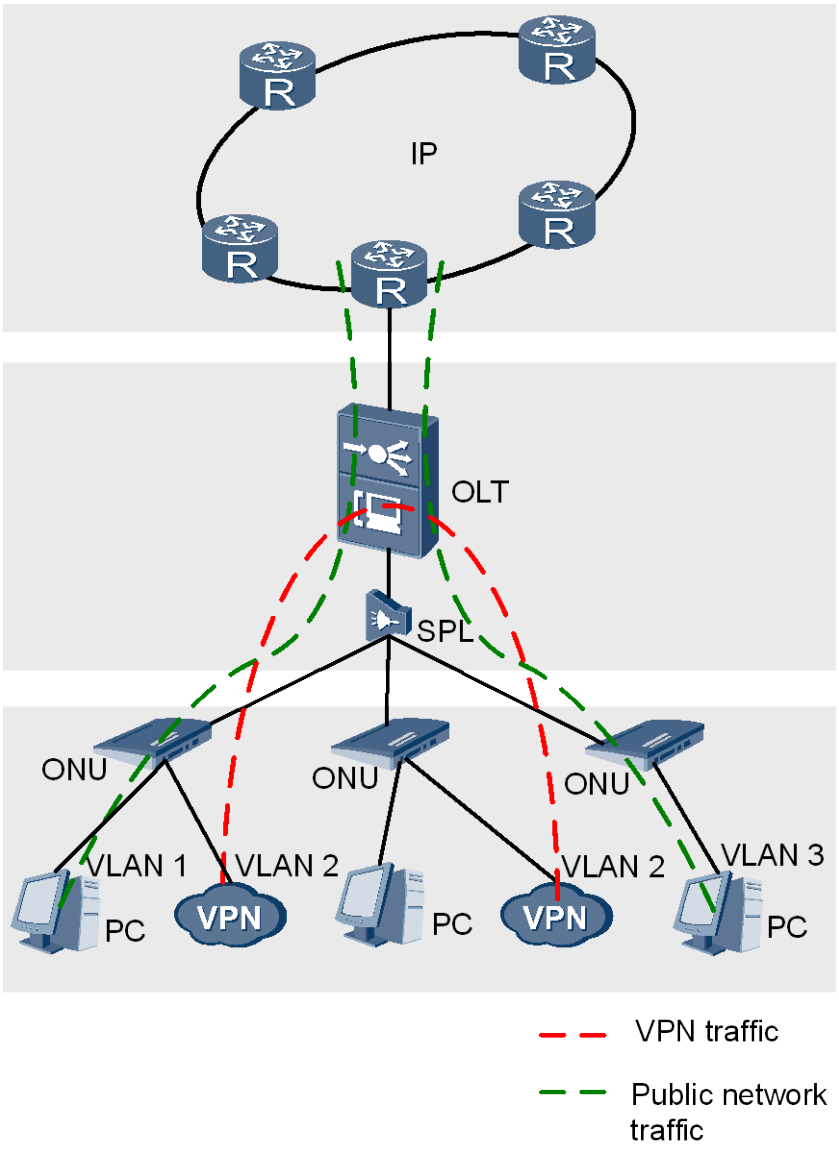
Figure 13-28 VLAN-based Layer 2 user bridging model



As shown in Figure 13-28, the following functions are implemented on the MA5600T/MA5603T/MA5608T after Layer 2 user bridging is enabled on a VLAN.

- Inter-board user bridging: bridging among ports on different boards, for example, users 1, 2, and 3 can exchange data with users 4 and 5
- Intra-board user bridging:
  - Bridging among different ONUs under the same PON port, for example, users 1 and 2 can exchange data with each other
  - Bridging among ONUs under different PON ports on the same PON board, for example, users 1, 2, and 3 can exchange data
  - Bridging among different ports on the same board, for example, users 4 and 5 can exchange data with each other

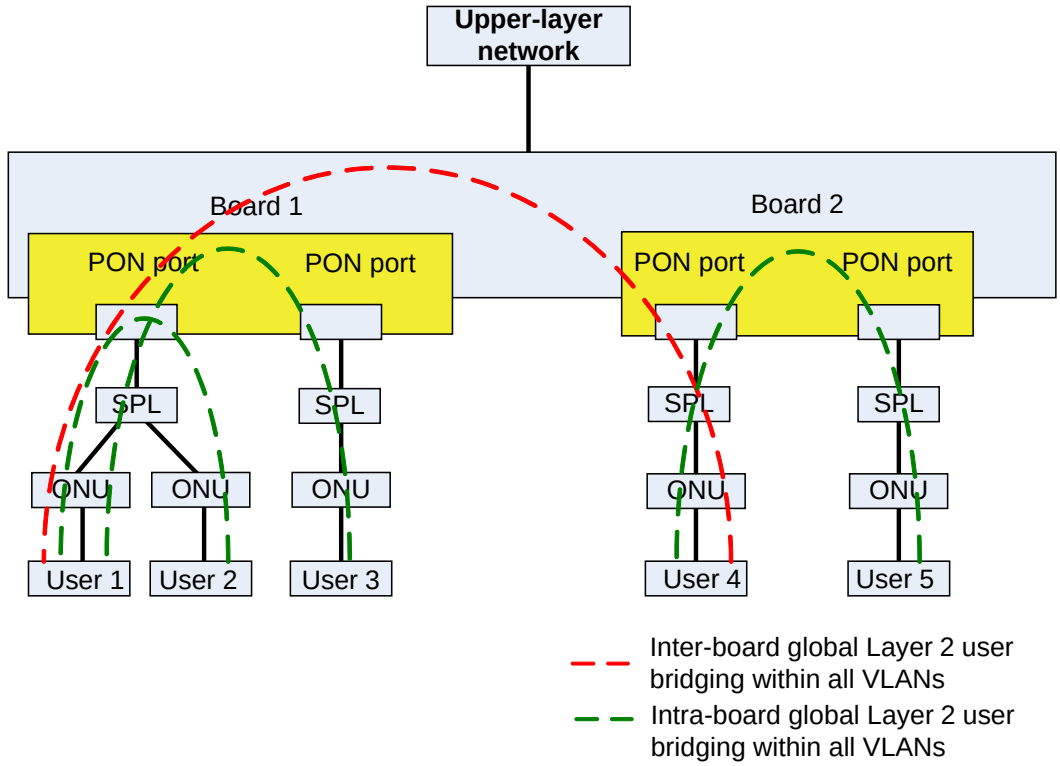
Figure 13-29 VLAN-based user bridging network



As shown in Figure 13-29, on one MA5600T/MA5603T/MA5608T, Layer 2 bridging is implemented among enterprise private line users (represented by the red line) but is not implemented among some common access users (represented by the green line). Services are classified by VLAN. VLAN 2 is a QinQ VLAN for carrying enterprise private line services. Enable Layer 2 user bridging on VLAN 2 so that private line users on the same MA5600T/MA5603T/MA5608T can exchange data.

Global Layer 2 User Bridging

Figure 13-30 Global Layer 2 user bridging model



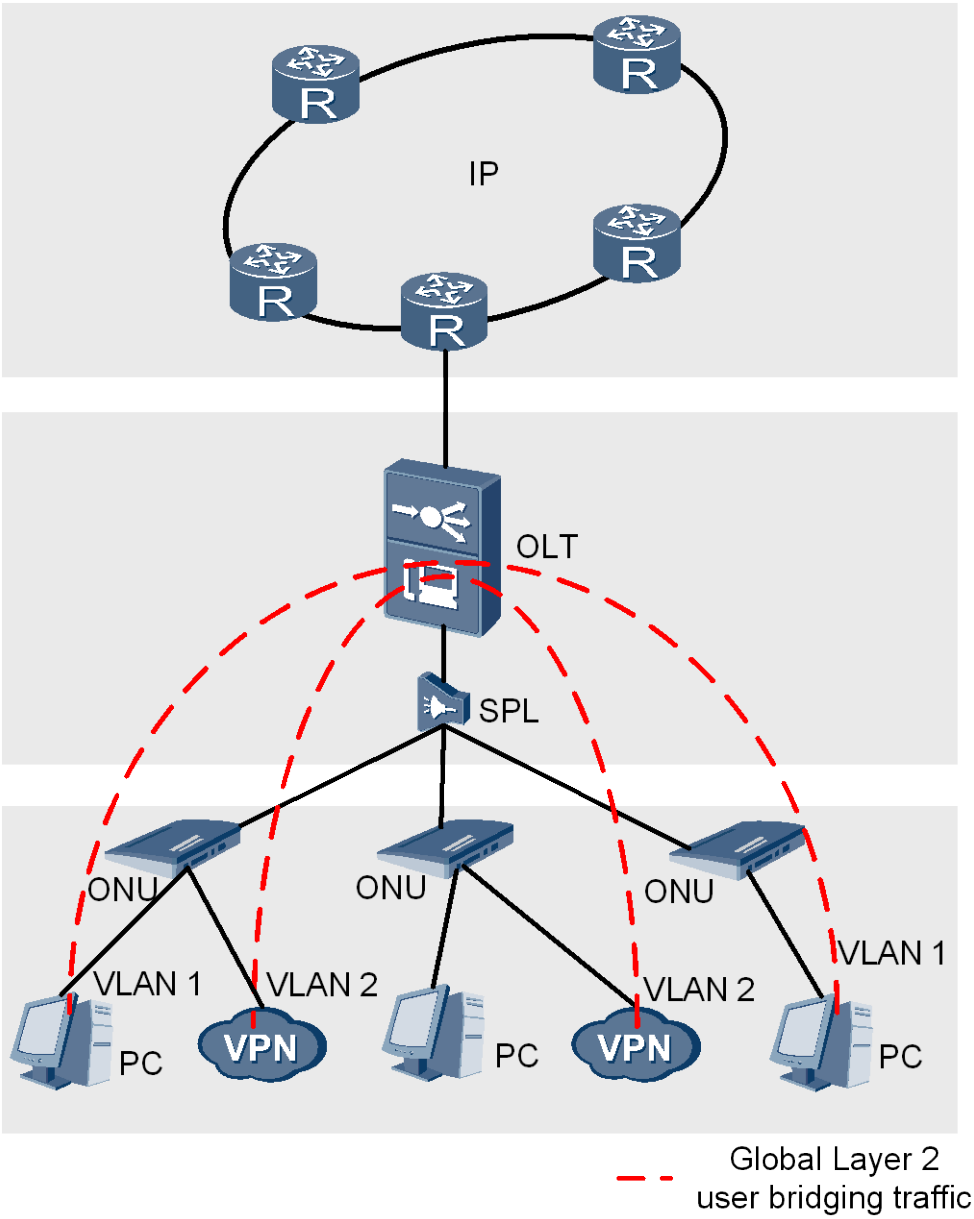
As shown in Figure 13-30, after global Layer 2 bridging is enabled for access users, Layer 2 bridging is enabled in all VLANs of the MA5600T/MA5603T/MA5608T.

NOTE

Layer 2 bridging takes effect for users only in the same VLAN but does not take effect for users in different VLANs.

- Inter-board user bridging: bridging among ports on different boards, for example, users 1, 2, and 3 can exchange data with users 4 and 5
- Intra-board user bridging:
  - Bridging among different ONUs under the same PON port, for example, users 1 and 2 can exchange data with each other
  - Bridging among ONUs under different PON ports on the same PON board, for example, users 1, 2, and 3 can exchange data

Figure 13-31 Global Layer 2 user bridging network



As shown in Figure 13-31, on the same MA5600T/MA5603T/MA5608T, all users in the same VLAN on the boards that support Layer 2 bridging can exchange data.

13.7.3 Configuration

This topic describes how to configure global and VLAN-based Layer 2 user bridging.

Configuring VLAN-based Layer 2 User Bridging

Default Setting

By default, Layer 2 user bridging based on VLAN service profiles is disabled.

Procedure

In global configuration mode, run the **vlan service-profile** command to enter the VLAN service profile mode.

Step 1 Run the **user-bridging** command to enable Layer 2 user bridging on a VLAN service profile.

Step 2 Run the **commit** command for the settings to take effect.

Step 3 Go back to the global configuration mode and run the **vlan bind service-profile** command to bind a VLAN to the VLAN service profile.

NOTE
If the VLAN has been bound to the VLAN service profile, you do not need to bind the VLAN again.

----End

Result

If you run the **display vlan service-profile** command, **User-bridging** is **enable** for the VLAN service profile. Then, under the MA5600T/MA5603T/MA5608T, users in the same VLAN with Layer 2 user bridging enabled can be pinged with each other

Example

Example: VLAN 10 is the S-VLAN for enterprise private line services. The VLAN is bound to VLAN service profile 3. Run the following commands to enable Layer 2 user bridging on VLAN 10 so that users in the VLAN on the same MA5600T/MA5603T/MA5608T can communicate with each other at Layer 2.

```
huawei(config)#vlan service-profile profile-id 3
huawei(config-vlan-srvprof-3)#user-bridging enable
huawei(config-vlan-srvprof-3)#commit
huawei(config-vlan-srvprof-3)#quit
```

Configuring Global Layer 2 User Bridging

Default Setting

By default, global Layer 2 user bridging is disabled in the system.

Procedure

In global configuration mode, run the **vlan-isolate disable** command to enable global Layer 2 user bridging.

----End

Result

If you run the **display vlan-isolate** command, Layer 2 isolation is **disable**. After the configuration, users in all VLANs on the boards that support Layer 2 user bridging on the same MA5600T/MA5603T/MA5608T can ping each other.

Example

Example: On the MA5600T/MA5603T/MA5608T, the main control board is SCUL and the service board is GPBC. Run the following command to enable global Layer 2 user bridging:

```
huawei(config)#vlan-isolate disable
  Command:
         vlan-isolate disable
  It will take several minutes, and console may be timeout, please use command
idle-timeout to set time limit
  Are you sure to switch the VLAN isolate? (y/n)[n]:y
```

13.7.4 Reference Standards and Protocols

The reference protocol of this feature is as follows:

DSL Forum TR-101: Migration to Ethernet-Based DSL Aggregation

14 QoS

About This Chapter

Quality of service (QoS) is a mechanism to ensure that user requirements for bandwidth, latency, jitter, and packet loss rate are met.

14.1 Introduction to QoS

Definition

Quality of service (QoS) is a mechanism that guarantees an expected service level with respect to bandwidth, latency, jitter, and packet loss in a communication network. The following indicators are used to measure QoS:

- Bandwidth: theoretical transmission capacities of a connection.
- Latency: the time required for information to travel from one network node to another network node. A high latency impacts the quality of real-time services (for example, IP telephone service).
- Jitter: variations in latency. Jitter can severely affect the quality of multimedia services (for example, VoD).
- Packet loss rate: the percentage of lost packets to the total packets during network transmission.

Purpose

QoS achieves the following purposes:

- Provides users with assured bandwidth.
- Regulates and controls IP network traffic.
- Reduces packet loss rate.
- Specifies packet priorities.
- Avoids and manages network congestion.
- Provides differentiated services for users.

14.2 QoS Models

QoS Models

The following table describes three types of QoS models that may be used in a network. The QoS model implemented on the MA5600T/MA5603T/MA5608T is the differentiated service (DiffServ) model. The following topics use the DiffServ model to describe QoS unless otherwise specified.

Table 14-1 Three QoS models

Type	Feature	Application
Best-effort service model	In this model, the network forwards data at best-effort rates and no guarantees are provided. The network drops data after all bandwidth is exhausted. This is a simple and unitary service model. It is the default service model for IP networks.	This service model applies to most data services, for example, Email services.
Integrated service model	This model is based on resource reservation. In this model, applications instruct a network to reserve bandwidth using the Resource Reservation Protocol (RSVP), and each unit in the network has to reserve bandwidth for specific data streams.	This service model is not widely used because of the following limitations: • End-to-end support is required for RSVP. • RSVP is not highly extendible. • RSVP protocol packets require large overhead.
DiffServ model	This model is based on priorities. In this model, the network identifies each data stream and provides corresponding QoS guarantees for the data streams. Traffic classification and priority marking are the prerequisites for using this model.	This model involves simple packet processing and is highly extendible. It is used for the following services: • VoD • Streaming media • VoIP • Video conferencing • Private-line services

QoS Components in the DiffServ Model

The DiffServ model uses four QoS components, which are described in the following table. You can design holistic QoS policies through flexible combinations of these components.

Table 14-2 QoS components in the DiffServ model

Component	Description
Traffic classification and priority marking	DiffServ operates on the principle of traffic classification and priority marking. Traffic classification: Data packets are placed into different traffic classes. Traffic classification does not modify the data packets. Priority marking: Data packets of each traffic class are marked with a specific priority. Priority marking modifies the data packets because the marking changes the values of some packet fields.
Traffic policing and shaping	Before providing services for a subscriber, service providers usually sign a service level agreement (SLA) with the subscriber to define the level of service. Traffic policing: Packets exceeding the SLA will be dropped. Traffic shaping: Packets exceeding the SLA are buffered and the transmission of these packets is resumed when bandwidth is sufficient.
Congestion management	Congestion management controls the sequence of data transmission when congestion occurs in a network.
Congestion avoidance	Congestion avoidance is a traffic control mechanism that actively drops packets when congestion occurs in a network. This mechanism addresses network overload by regulating the network traffic.

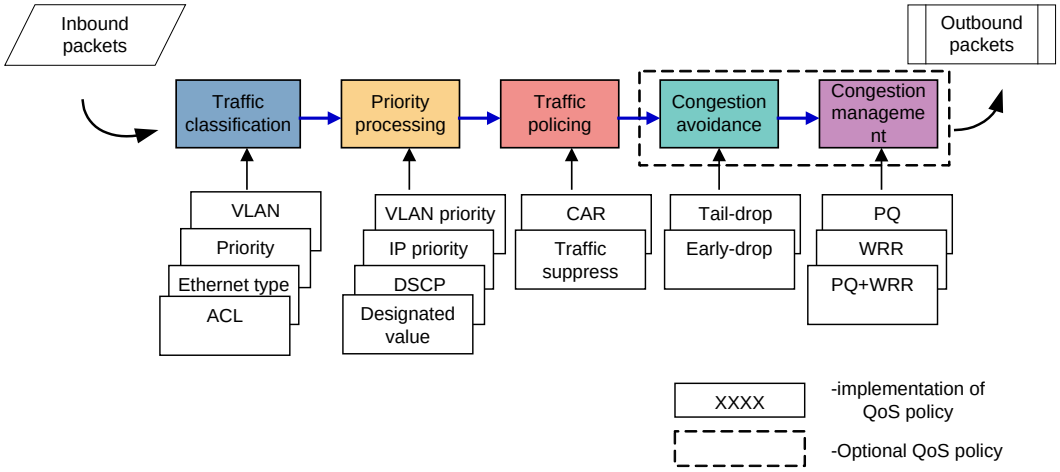
14.3 QoS Scheme

QoS Scheme

The MA5600T/MA5603T/MA5608T use the differentiated service (DiffServ) model. Figure 14-1 shows the QoS scheme of this model. The QoS scheme works as follows:

- Before forwarding the packets through an outbound interface, the MA5600T/MA5603T/MA5608T implements QoS policies, such as traffic classification, priority processing, traffic policing, congestion avoidance, and congestion management, on the packets.
- Congestion avoidance and congestion management are optional. They are not required if no congestion occurs (when the backplane bandwidth is higher than or equal to the maximum bandwidth actually required by a board).
- Each QoS policy can be implemented using multiple techniques. For details, see description about the QoS policies.

Figure 14-1 QoS scheme



QoS Congestion Point

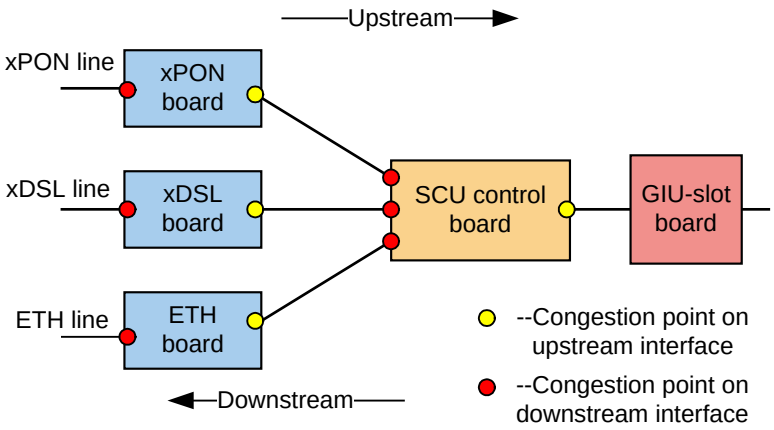
A network point where traffic congestion is likely to occur is defined as a QoS congestion point. Congestion avoidance and congestion management must be applied to such a point. As shown in Figure 14-2, there are QoS congestion points on both the control boards and service boards.

- There is a QoS congestion point on each service board in the downstream direction.
- In the upstream direction of a service board, congestion control is applied depending on the actual demand on bandwidth.

NOTE

- In the upstream direction, GPON boards do not have a congestion issue but 10G PON boards have.
- GIU-slot boards with ETH ports transparently transmit packets without QoS congestion control.

Figure 14-2 Upstream and downstream QoS congestion points



QoS Policies for Service Boards and Control Boards

The following table provides the QoS policies for service boards and control boards.

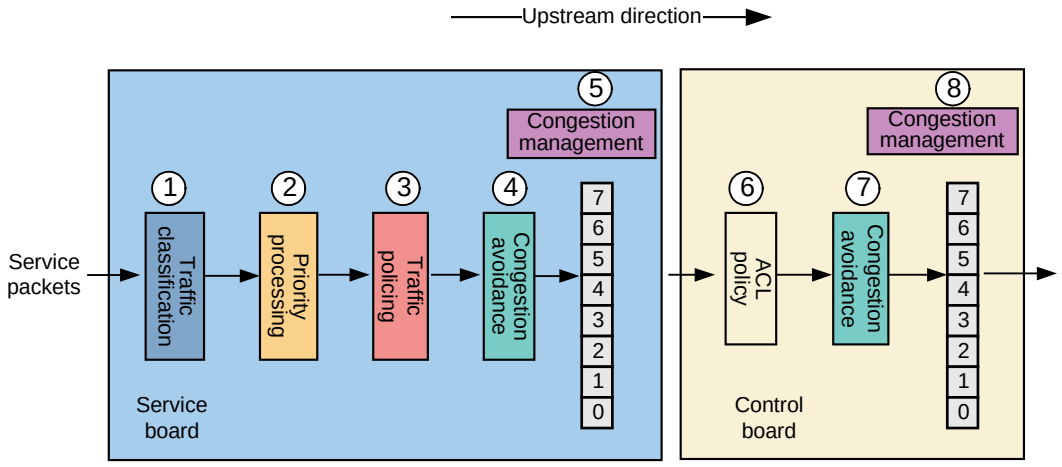
Board Type	QoS Policy	Remarks
Service board	Traffic classification Priority processing Traffic policing Congestion avoidance Congestion management	• Traffic classification on a service board is performed only in the upstream direction. • Congestion avoidance and congestion management are performed on an outbound interface of a service board.
Control board	Traffic classification (based on ACL) Priority processing (based on ACL) Traffic policing (based on ACL) Congestion avoidance Congestion management	Congestion avoidance and congestion management are performed on an outbound interface of a control board. For traffic classification, priority processing, and traffic policing based on ACL on the control board, see "14.10 ACL".

14.4 QoS Processing

General QoS Processing - Upstream Direction

Figure 14-3 shows the general QoS processing in the upstream direction.

Figure 14-3 QoS processing in the upstream direction



After the user packets enter a service board from a user port, the service board implements QoS processing on the packets. The following table describes the QoS processing.

Table 14-3 QoS processing on the service board

Step	QoS Policy	Example
1	Traffic classification: The service board places the Ethernet user packets into different classes based on packet characteristics and predefined QoS rules, and services the traffic differently.	Plan different VLANs or priorities for Internet, voice, and IPTV services to distinguish them, because these services have different QoS requirements.
2	Priority processing: When congestion occurs on the local device or the upper-layer network, the service board marks or remarks priorities for the user packets so that the packets are scheduled based on their priorities.	Assign priorities 0, 5, and 4 for Internet, voice, and IPTV services respectively.
3	Traffic policing: The service board implements traffic policing to limit the traffic and burst size of the user packets destined for a connection on the network. When the packets meet certain conditions, for example, when the traffic destined for the connection is too heavy, the service board takes actions accordingly, such as dropping or coloring (resetting the priorities) the packets. Traffic policing enables a port to work at a stable rate, preventing impact on lower-layer devices.	For Internet services, traffic policing provides a committed information rate (CIR) of 8 Mbit/s and limits the peak information rate (PIR) to 10 Mbit/s. When the rate of the Internet service packets is between the CIR and PIR, the service board marks the packets yellow. When the rate exceeds the PIR, the service board drops the packets.
4	Congestion avoidance: When congestion occurs during enqueuing on an outbound interface, the service board drops non-conformant packets to avoid further congestion.	To promptly drop packets based on priority, set the early drop threshold to 30% for 0-priority Internet services. After this setting takes effect, if 30% of the packets that arrive at a queue are Internet service packets, subsequent Internet service packets will be dropped.
5	Congestion management: When congestion occurs during dequeuing on an outbound interface, the service board uses queuing mechanisms to provide QoS guarantees to high-priority packets preferentially.	Use the strict priority (SP) scheduling algorithm to ensure that voice services with priority 5 are scheduled preferentially and Internet services with priority 0 are scheduled last when congestion occurs.

After the packets enter the control board, the control board implements QoS processing. The following table describes the QoS processing.

Table 14-4 QoS processing on the control board

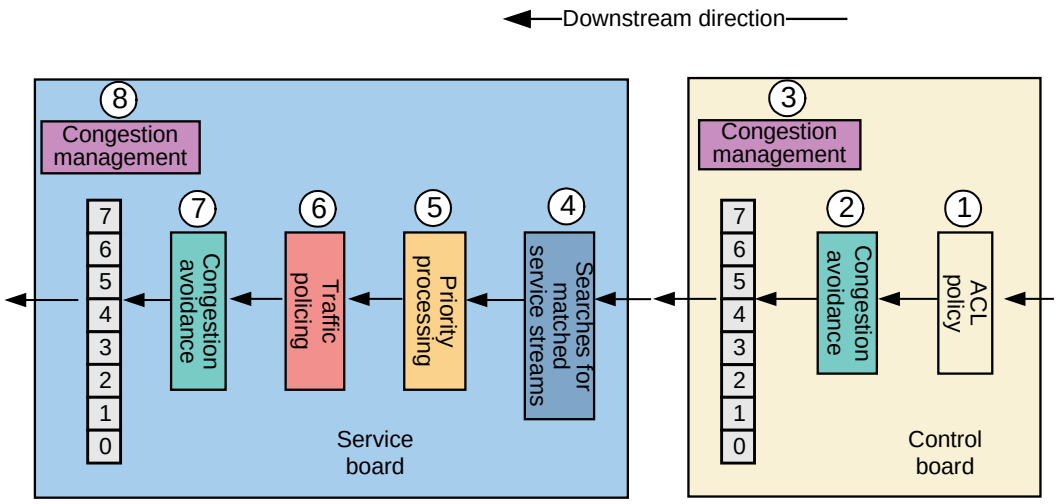
Step	QoS Policy	Example
6	ACL policy: The control board filters the data packets based on predefined ACL rules. After identifying data packets that match the ACL rules, the control board permits or denies the data packets based on the predefined policy.	Set ACL rules to ensure that only Internet and voice service packets carrying the matched VLAN tags are permitted to pass. Non-conformant service packets are dropped.
7	Congestion avoidance: When congestion occurs during enqueuing on an inbound interface, the control board drops non-conformant packets to avoid further congestion.	To promptly drop packets based on priority, set the early drop threshold to 30% for 0-priority Internet services. After this setting takes effect, if 30% of the packets that arrive at a queue are Internet service packets, subsequent Internet service packets will be dropped.
8	Congestion management: When congestion occurs during dequeuing on an outbound interface, the control board uses queuing mechanisms to provide QoS guarantees to high-priority packets preferentially.	Use the SP scheduling algorithm to ensure that voice services with priority 5 are scheduled preferentially and Internet services with priority 0 are scheduled last when congestion occurs.

General QoS Processing - Downstream Direction

Figure 14-4 shows QoS processing in the downstream direction. Different from QoS processing in the upstream direction, QoS processing in the downstream direction is performed in the following way:

- User packets are first processed by the control board and then by the service board.
- The service board does not perform traffic classification in the downstream direction. Instead, the service board searches for matched service streams based on a forwarding mode (VLAN+MAC or S-VLAN+C-VLAN).

Figure 14-4 QoS processing in the downstream direction



14.5 Traffic Classification

Traffic classification differentiates services by packet classification according to the characteristics of user Ethernet packets and certain rules, to implement different processing operations and provide different services.

14.5.1 Introduction

Traffic

Service flow or data streams are a type of traffic. A service flow is a set of packets that have common properties. For example, the Internet service packets of a user are one service flow and voice service packets of the user are another service flow.

In the MA5600T/MA5603T/MA5608T, a service flow is also called a service port.

For more details about the traffic, please refer to "13.4 Service Flow".

Traffic Classification

Traffic classification is a technique that categorizes user packets into different classes based on the properties of Ethernet packets and predefined QoS rules. Traffic classes are processed differently and this allows for the provision of differentiated services.

For example, to provide Internet, voice, and IPTV services simultaneously to the same user, the service packets must be separated into three classes.

Purpose

Traffic classification is used to support concurrent uses of multiple services. It differentiates service flows allowing predefined QoS guarantees to be provided for each of the service flows. The system performs service mappings based on service flows and performs QoS actions based on the service mappings. The QoS actions include switching between a user VLAN and

the network VLAN, upstream and downstream CAR policing, priority marking, and queue scheduling.

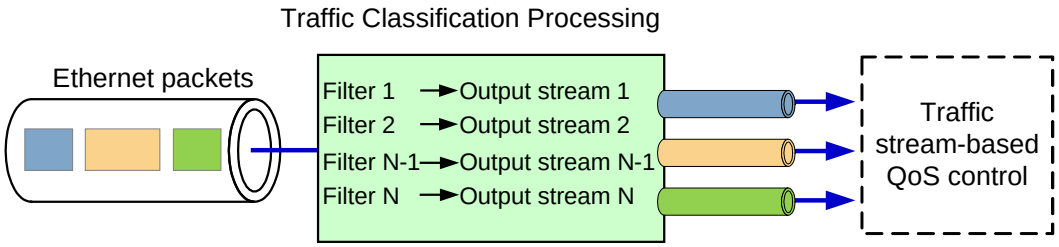
14.5.2 Implementation Principle

Traffic Classification Process

Figure 14-5 shows the traffic classification process. Traffic classification is performed using classifiers. Each classifier includes a filter (one or more classification rules as described in Table 14-5) and an output flow. The classifiers separate Ethernet packets received by a device into different traffic classes. QoS guarantees are then provided for each of the traffic classes.

Traffic classification is based on Ethernet packets. For ATM ports in xDSL and PON access mode, the system segments and reassembles incoming service packets to recover Ethernet frames. The system then performs traffic classification for the Ethernet frames. Therefore, each VC port, GEM port, or EPON logical link identifier (LLID) can be regraded as a logical port of an Ethernet port. Traffic classification is performed based on this logical port and this logical port receives and sends Ethernet frames.

Figure 14-5 Traffic classification process



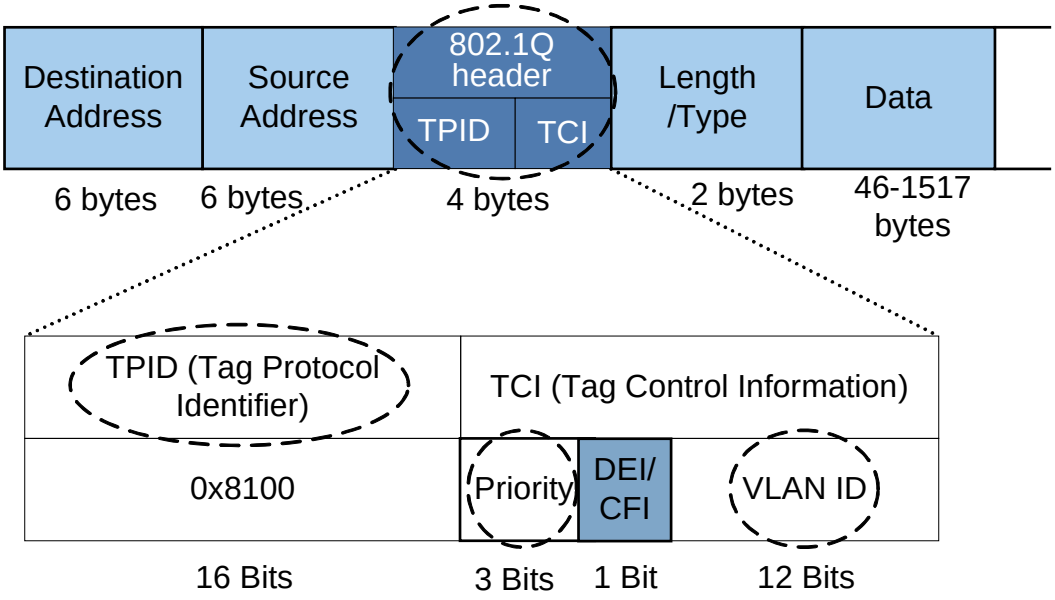
The system performs traffic classification only for upstream packets. Meanwhile, the system learns the MAC addresses of the packets when the packets are forwarded based on VLAN+MAC forwarding mode, and records the indexes of service flows. In the downstream direction, the system searches for the service flow based on the VLAN and MAC address.

Traffic Classification Method and Application

Traffic classification is performed to distinguish user services based on the properties of Ethernet packets. Traffic classification is usually performed based on the following fields in Ethernet frame headers (shown in the Figure 14-6):

- VLAN
- Priority
- Ethernet type

Figure 14-6 Ethernet frame format defined by 802.1q



For detailed traffic classification rules, see Table 14-5.

Table 14-5 Traffic classification rules

Traffic Classification Rule	Application Scenario
Based on VLAN	If services have different C-VLANs, traffic classification can be performed based on the C-VLANs. For example, traffic classification can be based on the C-VLAN of 100 for Internet services and 200 for voice services.
Based on C-VLAN+Ethernet type	If a single field cannot be used to distinguish the service type of packets, traffic classification can be based on combinations of the three fields in Ethernet packets. For example, if services have the same VLANs and priorities, the C-VLAN+Ethernet type combination can be used to distinguish the services.
Based on C-VLAN+priority	
Based on S-VLAN+C-VLAN	This rule is used when users of a service type must be identified uniquely. For example, if a C-VLAN is allocated to a specific user, the S-VLAN+C-VLAN combination can be used to identify this user.
Based on other-all	When a common service and a transparent LAN service exist in a system, other-all can be used to distinguish the two services. For example, a system has an Internet service (VLAN 100) and other transparent LAN services. If the Internet service must be distinguished from the other transparent LAN services, the transparent LAN services can be mapped into the other-all traffic class.

Note:

- S-VLAN refers to service VLAN and is usually used to identify a service. In a service flow, S-VLAN is also called network-side VLAN or outer VLAN.
- C-VLAN refers to customer VLAN and is usually used to identify a user. In a service flow, C-VLAN is also called inner VLAN.

Generally, one logical port supports only one traffic classification rule, excluding the following two situations:

- The single-tagged VLAN-based (**user-vlan** { **untagged** | *user-vlanid* }) traffic classification and the double-tagged VLAN-based (**double-vlan outer-vlan** *vlanid* **inner-vlan** *vlanid*) traffic classification can coexist, which is supported by only SPUA and OPGD.
- The single-service-based traffic classification and multi-service-based traffic classification can coexist on the port whose PVC encapsulation type is auto-sensing (**encapsulation type auto**). The system automatically matches the single-service flows or multi-service flows according to the learned PVC ATM adaptation layer type 5 (AAL5). Then, the single-service flows are used for IPoA/PPPoA services and multi-service flows are used for the Ethernet service.

Packet Matching Priority

- Packets are matched first to the rules that define specific classifiers (for example, C-VLAN or priority) and, if they cannot be matched to a specific traffic class, they are matched to the other-all class.
- If no other-all traffic class has been configured, all incoming packets are matched with specific traffic classes.
- If incoming packets cannot be matched with any traffic class, they are dropped.

14.5.3 Configuring the Traffic Classification

Context

A service flow is the result of traffic classification based on physical ports or logical ports. Physical ports that can be used for traffic classification include Ethernet ports and VDSL ports in PTM mode. Logical ports include xDSL ports in ATM mode, VC ports in ATM access mode, GEM ports in GPON access mode, and LLID ports in EPON access mode.

Procedure

Run the **service-port** command to perform traffic classification.

```
(config)#service-port
{ desc<K>|index<U><0,32767>|remote-desc<K>|source<K>|uplink-port<K>|vlan<K> }:vlan
{ aoe<K>|vlanid<U><1,4093> }:100
{ adsl<K>|atm<K>|epon<K>|eth<K>|gpon<K>|port<K>|shdsl<K>|vdsl<K> }:eth
{ frameid/slotid/portid<S><Length 1-15> }:0/3/0    //Indicates the logical port of the
service flow.
{ <cr>|bundle<K>|inbound<K>|multi-service<K>|rx-cttr<K>|tag-transform<K> }:multi
-service     //Indicates the multi-service mode. In this mode, a logical port carries
multiple services.
{ double-vlan<K>|user-8021p<K>|user-encap<K>|user-vlan<K> }:user-vlan   //Indicate
```

```
various modes of traffic classification.
{ other-all<K>|priority-tagged<K>|untagged<K>|user-vlanid<U><1,4095> }:
```

The key parameters related to traffic classification for this command are as follows:

- **multi-service**: indicates multiple services. Traffic classification is required when a service port carries multiple services.
- **double-vlan**: Traffic classification is performed based on S-VLAN+C-VLAN.
- **user-8021p**: Traffic classification is performed based on the user-side 802.1p priority.
- **user-encap**: Traffic classification is performed based on the user-side encapsulation type, IPoE or PPPoE.
- **user-vlan**: Traffic classification is performed based on the user-side VLAN. Valid values for the user-side VLAN are 1-4095, untagged, priority-tagged (the user packet is tagged as VLAN 0), and other-all.

NOTE
For details on this command, see "Service Virtual Port Configuration" in Command Reference.

Step 1 Run the **display service-port** command to verify that the traffic classification has been applied to the service port.

----End

Example

- Traffic classification based on C-VLAN

  To perform traffic classification for Ethernet port 0/3/0 based on C-VLAN 100, do as follows:

  ```
  huawei(config)#service-port vlan 8 eth 0/3/0 multi-service user-vlan 100
  ```

- Traffic classification based on priority

  To perform traffic classification for Ethernet port 0/3/0 based on priority 3, do as follows:

  ```
  huawei(config)#service-port vlan 8 eth 0/3/0 multi-service user-8021p 3
  ```

- Traffic classification based on Ethernet type

  To perform traffic classification for Ethernet port 0/3/0 based on Ethernet type PPPoE, do as follows:

  ```
  huawei(config)#service-port vlan 8 eth 0/3/0 multi-service user-encap pppoe
  ```

- Traffic classification based on C-VLAN+Ethernet type

  To perform traffic classification for Ethernet port 0/3/0 based on C-VLAN 100 and Ethernet type PPPoE, do as follows:

  ```
  huawei(config)#service-port vlan 8 eth 0/3/0 multi-service user-vlan 100 user-encap pppoe
  ```

- Traffic classification based on C-VLAN+priority

  To perform traffic classification for Ethernet port 0/3/0 based on C-VLAN 100 and priority 3, do as follows:

  ```
  huawei(config)#service-port vlan 8 eth 0/3/0 multi-service user-vlan 100 user-8021p 3
  ```

- Traffic classification based on C-VLAN+S-VLAN

To perform traffic classification for Ethernet port 0/3/0 based on S-VLAN 100 and C-VLAN 10, do as follows:

```
huawei(config)#service-port vlan 8 eth 0/3/0 multi-service double-vlan outer-vlan 100 inner-vlan 10
```

14.6 Priority Marking

According to different priority marking policies, the inner and outer VLAN priorities can be set for service-ports, or the user-side priority can be copied for service-ports.

14.6.1 Introduction

Definition

Priority processing is a process of marking or re-marking the priority for a packet so that equipment or network can process the packet based on the defined priority. This process is performed in the following way:

- Priorities of packets are usually marked at an inbound interface of equipment or a network and is re-marked inside the equipment or network.
- Packet priorities include both forwarding and drop priorities. These two priorities form the basis for QoS processing. With respect to drop priority processing, the equipment or network uses the drop eligibility indicator (DEI) in an Ethernet packet to mark the color of the packet.

Purpose

Priority processing is the basis for equipment or a network to schedule packets. When congestion occurs, equipment or a network schedules packets based on priorities.

14.6.2 Basic Concepts

Priorities that equipment processes mainly include the VLAN priority (802.1p priority) and IP precedence.

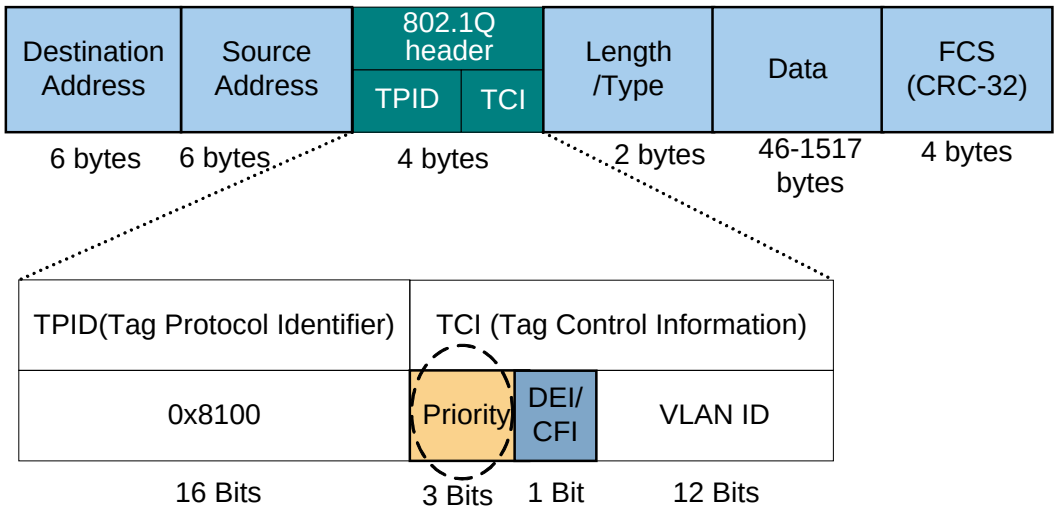
VLAN Priority

VLAN priority, also called 802.1p priority, refers to the packet priority defined at the link layer. This priority represents the class of service (CoS). As defined in IEEE 802.1q, the VLAN priority uses three bits in the VLAN tag (Figure 14-7 shows the position of the VLAN priority in an Ethernet frame).

As shown in the figure, the Priority field indicates the VLAN priority. This field consists of three bits. The value of the three bits ranges from 0 to 7. The value 0 is the lowest priority and 7 the highest. The priority values set in these fields determine the order in which packets are transmitted when congestion occurs on a port.

The DEI field consists of one bit and represents the drop eligible indicator defined in the 802.1ad protocol. It is used to color the packet. For example, the value 0 means green and 1 means yellow. When the function of color-based early drop is enabled, yellow packets are dropped with preference in case of congestion.

Figure 14-7 Ethernet frame format defined by 802.1q



IP Precedence

The IP protocol defines differentiated services code point (DSCP) and type of service (ToS). They occupy the same field (one byte) in an IP header. IP bearer network devices schedule and forward packets based on the DSCP or ToS filled to provide QoS guarantees for different services.

ToS in the IP header specifies a traffic class for a packet rather than a priority (which is determined by the device). The ToS field consists of eight bits, including a 3-bit IP precedence sub-field, 4-bit ToS sub-field, and one reserved bit (set to 0). The four bits in the ToS sub-field represent the minimum latency, maximum throughput, maximum reliability, and minimum cost. Only one of the four bits can be set to 1. If all the four bits are set to 0, the service is a common service.

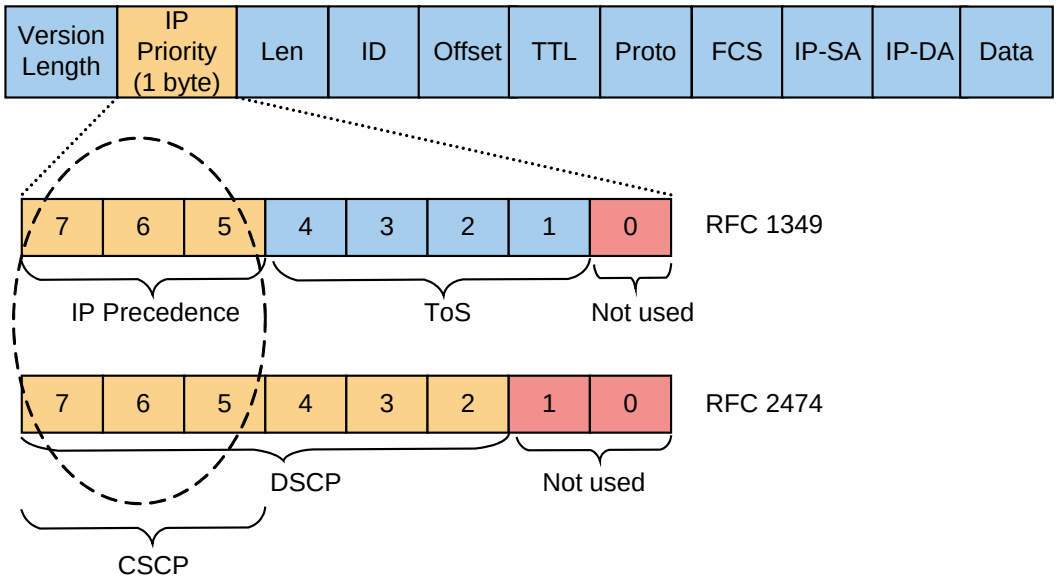
DSCP is defined in the RFC 2474, it is a re-defined object based on an IPv4 type of service (ToS) and an IPv6 traffic class.

NOTE

The traffic class (TC) field in the IPv6 packet header has the same functions as the IP priority field in the IPv4 packet header. This topic uses the IP priority field in the IPv4 packet header as an example to describe the IP priority.

As shown in Figure 14-8, the six most significant bits in the DS field (bits 7-2) are used as the DS CodePoint (DSCP) and the two least significant bits (bits 1 and 0) are reserved. The three most significant bits in the DS field (bits 7-6-5) are the class selector code point (CSCP), which indicates a DSCP type.

Figure 14-8 IPv4 packet format



DSCP is used to select per-hop behavior (PHB) on each node of a network. PHBs describe the external behaviors that are visible when the DS node is used for data stream aggregation. IETF has defined the following types of PHBs:

- Class selector (CS)
- Expedited forwarding (EF)
- Assured forwarding (AF)
- Best-effort (BE)

Table 14-6 shows the structure of PHBs.

Table 14-6 Structure of PHBs

PHB	Bit 7-6-5	Bit 4-3-2
CS	aaa (Remark)	000
BE	000	000
EF	101	110
AF	bbb (Remark)	cc0 (Remark)

Remark: a, b and c indicates a single bit, whose value is 0 or 1. Where,

- The "aaa" has eight values, from 000 to 111, it corresponds to the decimal number 0-7. And it can map with IP precedence individually.
- The "bbb" has four values: 001, 010,011 and 100, it corresponds to the decimal number 1-4.
- The "cc" has three values, from 01 to 11, it corresponds to the decimal number 1-3.

Table 14-7 shows the common DSCP service types and corresponding priorities.

Table 14-7 Common DSCP service types and corresponding priorities

Service type	IP precedence/MPLS EXP /802.1P priority	DSCP value (binary)	Application
BE	0	0	Internet
AF1	1	001 010	Leased Line
AF1	1	001 100	Leased Line
AF1	1	001 110	Leased Line
AF2	2	010 010	IPTV VoD
AF2	2	010 100	IPTV VoD
AF2	2	010 110	IPTV VoD
AF3	3	011 010	IPTV Broadcast
AF3	3	011 100	IPTV Broadcast
AF3	3	011 110	IPTV Broadcast
AF4	4	100 010	NGN/3G Singaling
AF4	4	100 100	NGN/3G Singaling
AF4	4	100 110	NGN/3G Singaling
EF	5	101 110	NGN/3G voice
CS6	6	110 000	Protocol
CS7	7	111 000	Protocol

14.6.3 Priority Sources

Priority processing includes copying, designating, and mapping inner and outer VLAN priorities.

Priority Processing for an Outer VLAN (or Single-Tagged VLAN) of Ethernet Service Flows

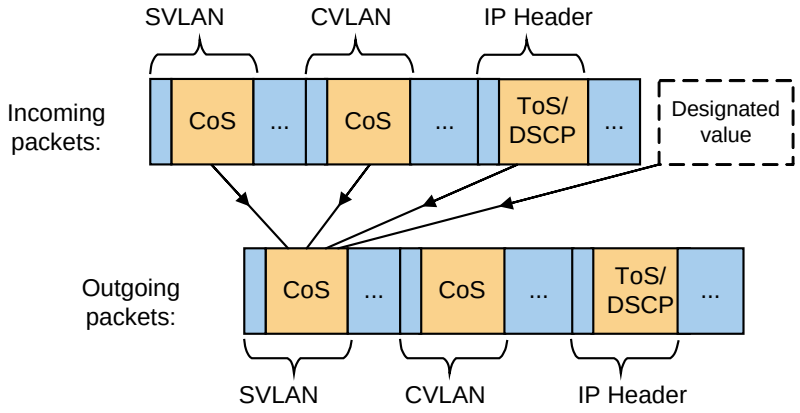
Figure 14-9 shows how the priority of an outer VLAN or the single-tagged VLAN is processed. The priority of an outer VLAN can be derived from multiple sources:

- Copied from the outer VLAN priority (user-cos) of an incoming packet
- Copied from the inner VLAN priority (user-inner-cos) of an incoming packet
- Copied or mapped from the IP ToS priority (user-tos) of an incoming packet
- Mapped from the IP DSCP priority (user-dscp) of an incoming packet

- Copied from the priority of a designated packet (prival)

NOTE
Currently, the downstream equipment of a service flow cannot copy the user-dscp priority.

Figure 14-9 Priority processing for an outer VLAN (or single-tagged VLAN) of Ethernet Service Flows



Some boards support several priority mapping choices:

- 802.1p priority -> 802.1p priority (To implement this mapping, run the **pbits-to-pbits mapping table** command. It's only valid for inner VLAN priority)
- IP priority -> 802.1p priority (To implement this mapping, run the **ipprec-to-pbits mapping table** command.)
- DSCP priority -> 802.1p priority (To implement this mapping, run the **dscp-to-pbits mapping table** command.)

Figure 14-10 shows the mappings between incoming and outgoing priorities of packets.

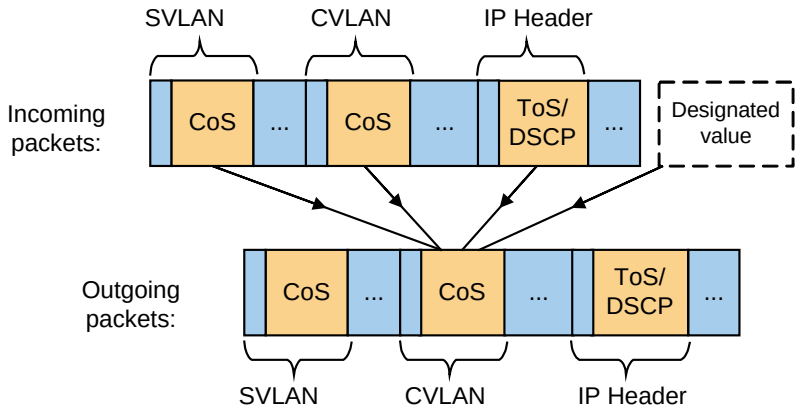
Figure 14-10 Mapping between the incoming and outgoing priorities of packets

802.1p/IP priority (Incoming packets)	DSCP priority (Incoming packets)	802.1p priority of outgoing packets (Copy priority by default, and priority can't be modified)	802.1p priority of outgoing packets (Mapping, priority can be modified)
0	0-7	0	0
1	8-15	1	1
2	16-23	2	2
3	24-31	3	3
4	32-39	4	3
5	40-47	5	3
6	48-55	6	4
7	56-63	7	5

Priority Processing for an Inner VLAN of Ethernet Service Flows

Figure 14-11 shows how the priority of an inner VLAN is processed. The sources of an inner VLAN priority are the same as those of an outer VLAN priority.

Figure 14-11 Priority processing for an inner VLAN

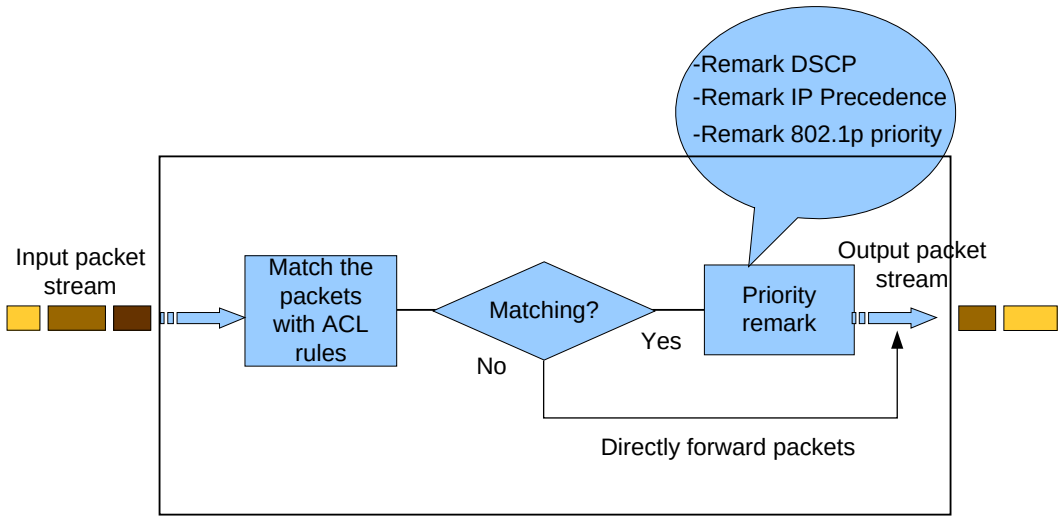


Priority Processing for ACL-based Packets

Figure 14-12 shows the policy of priority processing for ACL-based packets. The system processes the input packets according to ACL rules. After packets match ACL rules, the system remarks these packets by running the **traffic-priority** command.

- Specifies the DSCP priority of the packets.
- Specifies or copies the 802.1p priority of packets as the IP precedence priority of packets.
- Specifies or copies the IP precedence priority of packets as the 802.1p priority of packets.

Figure 14-12 Priority processing for ACL-based packets



14.6.4 Implementation Principle

Priority processing applies to common data packets and specific protocol packets. The priority processing mechanisms for data packets and protocol packets are different, because protocol packets are captured and processed by the CPU.

Priority processing mechanisms for common (non-service-bundle) and bundled (service-bundle) service flows are also different.

In addition, priority processing is affected by the following factors:

- Attribute parameters of service flows (whether user-802.1p is configured)
- 14.6.3 Priority Sources
- VLAN priority
- Packet forwarding mode (Layer 2 or Layer 3 forwarding)

Figure 14-13 describes packet priority processing in the non-service-bundle scenario.

Figure 14-13 Packet priority processing in the non-service-bundle scenario

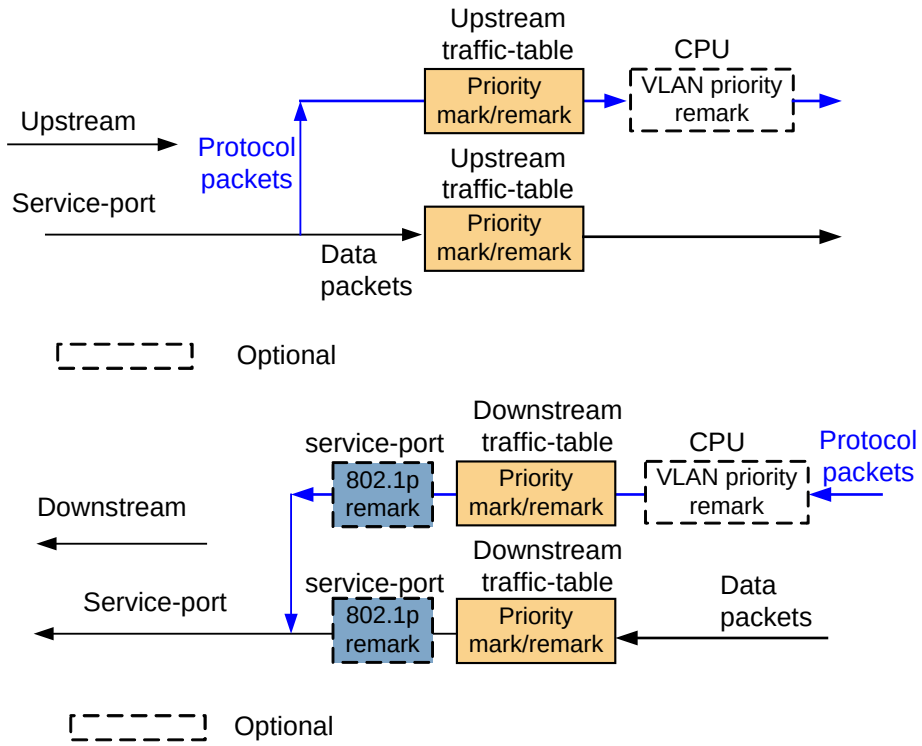
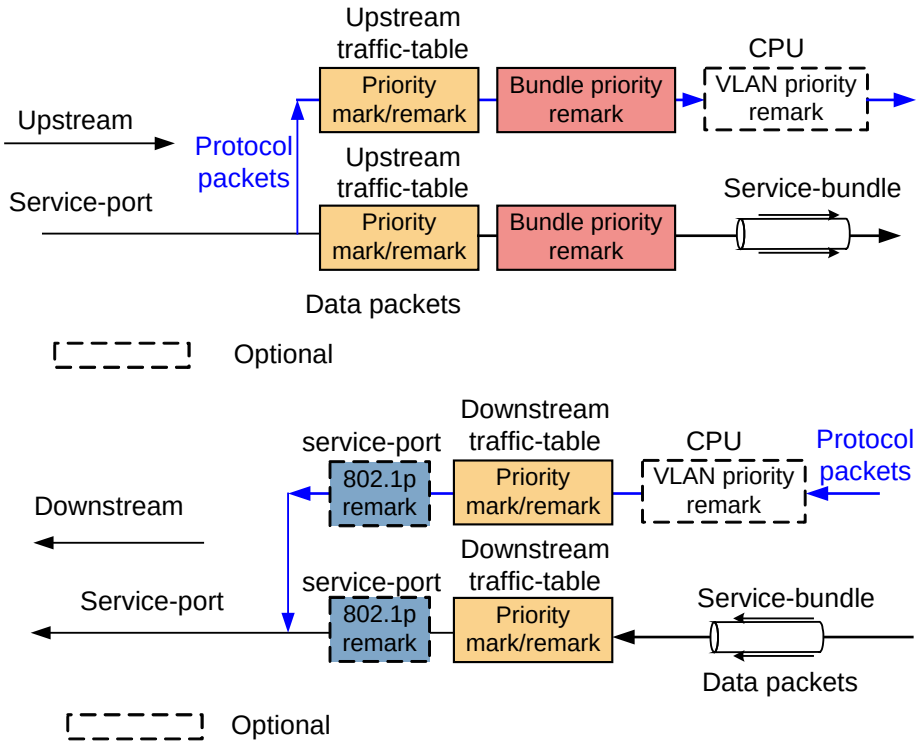


Figure 14-14 describes packet priority processing in the service-bundle scenario.

Figure 14-14 Packet priority processing in the service-bundle scenario



Note that the packet priority processing described in the preceding scenarios is a general processing model. Not all types of packets strictly comply with this model. For details on priority processing for packets of a particular type, refer to the priority processing mechanism for the packet type.

Priority Processing for Data Packets

This topic describes the priority processing for data packets. For detailed processing rules and results, refer to Priority Processing Rules.

Priority Processing for Data Packets in the Non-Service-Bundle Scenario

Priority processing for data packets in the non-service-bundle scenario has the following characteristics:

- The service flow priority is irrelevant to the VLAN priority (Run the **vlan priority** *vlanid priority* command to set VLAN priority).
- Priority processing in the upstream direction is based on the priority processing policy (specify or copy/map priorities, as shown in Figure 14-15) configured in the upstream traffic profile.
- If the **user-cos**, **user-tos**, or **user-inner-cos** parameter is selected but no such priority exists in the packet (in other words, priority copying fails), the system uses the default 802.1p priority for the packet.
- In the downstream direction, if the 802.1p priority (user-802.1p) is configured for the service flow, remarking is implemented based on the 802.1p priority; if not, priority processing is implemented based on the settings in the downstream traffic profile, as shown in Figure 14-16.

Priority processing in the downstream direction may have the following exception cases:

- If the service flow is created on the GPON/OPGD/OPGE board and the priority processing policy specified in the traffic profile in the downstream direction is **user-cos**, the 802.1p priority remarking does not take effect.
- If the service flow is created on the GPON/OPGD/OPGE board, the priority processing policy specified in the traffic profile in the downstream direction is **user-tos**, and the packet is an IP packet, the 802.1p priority remarking does not take effect.

Figure 14-15 Priority processing for upstream data packets in the non-service-bundle scenario

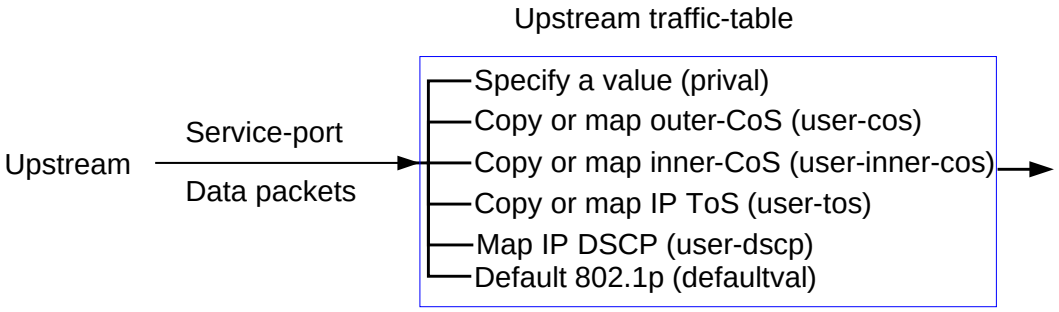


Figure 14-16 Priority processing for downstream data packets in the non-service-bundle scenario

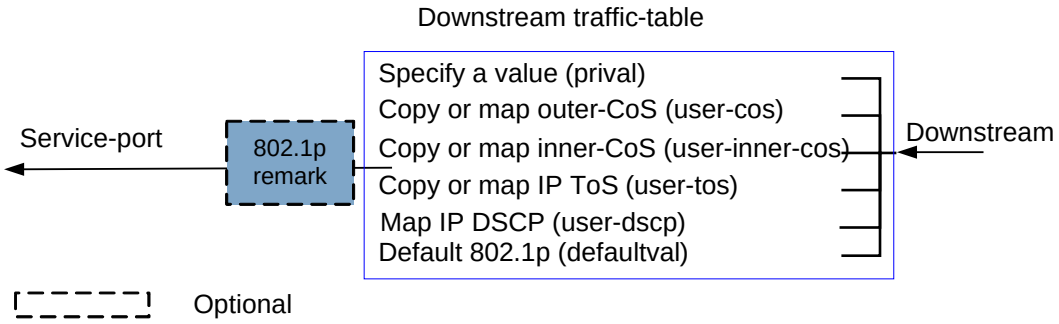


Figure 14-17 shows the detailed processing results.

Figure 14-17 Priority processing results of upstream and downstream data packets in the non-service-bundle scenario

service-port		traffic table				GPON & OPGD & OPGE System Behavior		xDSL & OPFA System Behavior	
user-8021p	user-vlan	priority	user-cos	user-tos	deafult	Upstream direction	Downstream direction	Upstream direction	Downstream direction
-	√	√	-	-	-	traffic table priority	traffic table priority	traffic table priority	traffic table priority
√	√	√	-	-		traffic table priority	user-8021p	traffic table priority	user-8021p
-	√	-	√	-	√	user-cos	network-cos	user-cos	network-cos
√	√	-	√	-	√	user-cos	network-cos	user-cos	user-8021p
-	√	-	-	√	√	user-tos	network-tos	user-tos	default priority
√	√	-	-	√	√	user-tos	network-tos	user-tos	user-8021p
√: Configured -:Unconfigured									

Priority Processing for Data Packets in the Service-Bundle Scenario

Priority processing for data packets in the service-bundle scenario has the following characteristics:

- The priority of upstream packets is the priority specified for the service-bundle group. The settings in the traffic profile do not apply.
- The priority processing for downstream packets is the same as that in the non-service-bundle scenario.

  Priority processing in the downstream direction may have the following exception cases:
  - If the service flow is created on the GPON/OPGD/OPGE board and the priority processing policy specified in the traffic profile in the downstream direction is **user-cos**, the 802.1p priority remarking does not take effect.
  - If the service flow is created on the GPON/OPGD/OPGE board, the priority processing policy specified in the traffic profile in the downstream direction is **user-tos**, and the packet is an IP packet, the 802.1p priority remarking does not take effect.

Figure 14-18 describes the priority processing for upstream data packets in the service-bundle scenario.

Figure 14-18 Priority processing for upstream data packets in the service-bundle scenario

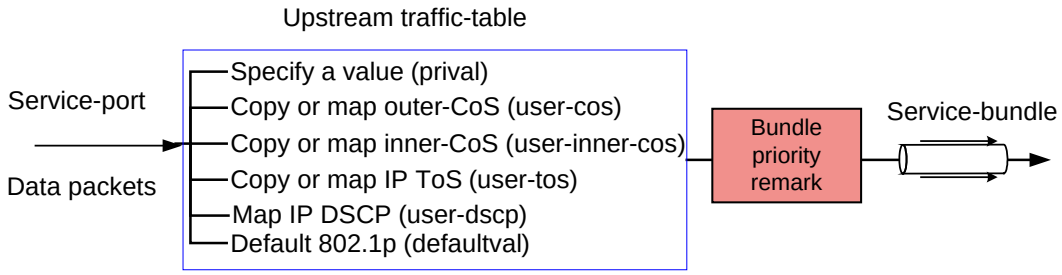


Figure 14-19 shows the detailed processing results.

Figure 14-19 Priority processing for upstream and downstream data packets in the service-bundle scenario

service-port		traffic table				GPON & OPGD & OPGE System Behavior		xDSL & OPFA System Behavior	
user-8021p	user-vlan	priority	user-cos	user-tos	deafult	Upstream direction	Downstream direction	Upstream direction	Downstream direction
-	√	√	-	-		bundle cos	traffic table priorty	bundle cos	traffic table priorty
√	√	√	-	-		bundle cos	user-8021p	bundle cos	user-8021p
-	√	-	√	-	√	bundle cos	network-cos	bundle cos	network-cos
√	√	-	√	-	√	bundle cos	network-cos	bundle cos	user-8021p
-	√	-	-	√	√	bundle cos	network-tos	bundle cos	default priorty
√	√	-	-	√	√	bundle cos	network-tos	bundle cos	user-8021p
√: Configured		-:Unconfigured							

Priority Processing for DHCP & PPPoE Packets

This topic describes the priority processing for DHCP & PPPoE packets. For detailed processing rules and results, refer to Priority Processing Rules.

Priority Processing for DHCP & PPPoE Packets in the Non-Service-Bundle Scenario

DHCP protocol packets can work in Layer 2 or Layer 3 relay mode. The DHCP protocol packets in this document refer to DHCP Option82 packets in Layer 2 relay mode and DHCP Layer 3 relay packets.

PPPoE protocol packets work only in Layer 2 mode. The PPPoE packets in this document refer to PPPoE PITP packets.

When DHCP and PPPoE packets work in Layer 2 mode, the priority processing mechanism is the same as that for common data packets, which has the following characteristics:

- The service flow priority is irrelevant to the VLAN priority.
- Priority processing in the upstream direction is based on the priority processing policy (specify or copy/map priorities) configured in the upstream traffic profile.
- If the **user-cos**, **user-tos**, or **user-inner-cos** parameter is selected but no such priority exists in the packet (in other words, priority copying fails), the system uses the default 802.1p priority for the packet.

NOTE

When the user-tos priority policy is specified in the traffic profile, ToS copying fails for PPPoE PITP packets. When this occurs, the default 802.1p priority in the traffic profile will be used.

- In the downstream direction, if the 802.1p priority (user-802.1p) is configured for the service flow, remarking is implemented based on the 802.1p priority; if not, priority processing is implemented based on the settings in the downstream traffic profile.

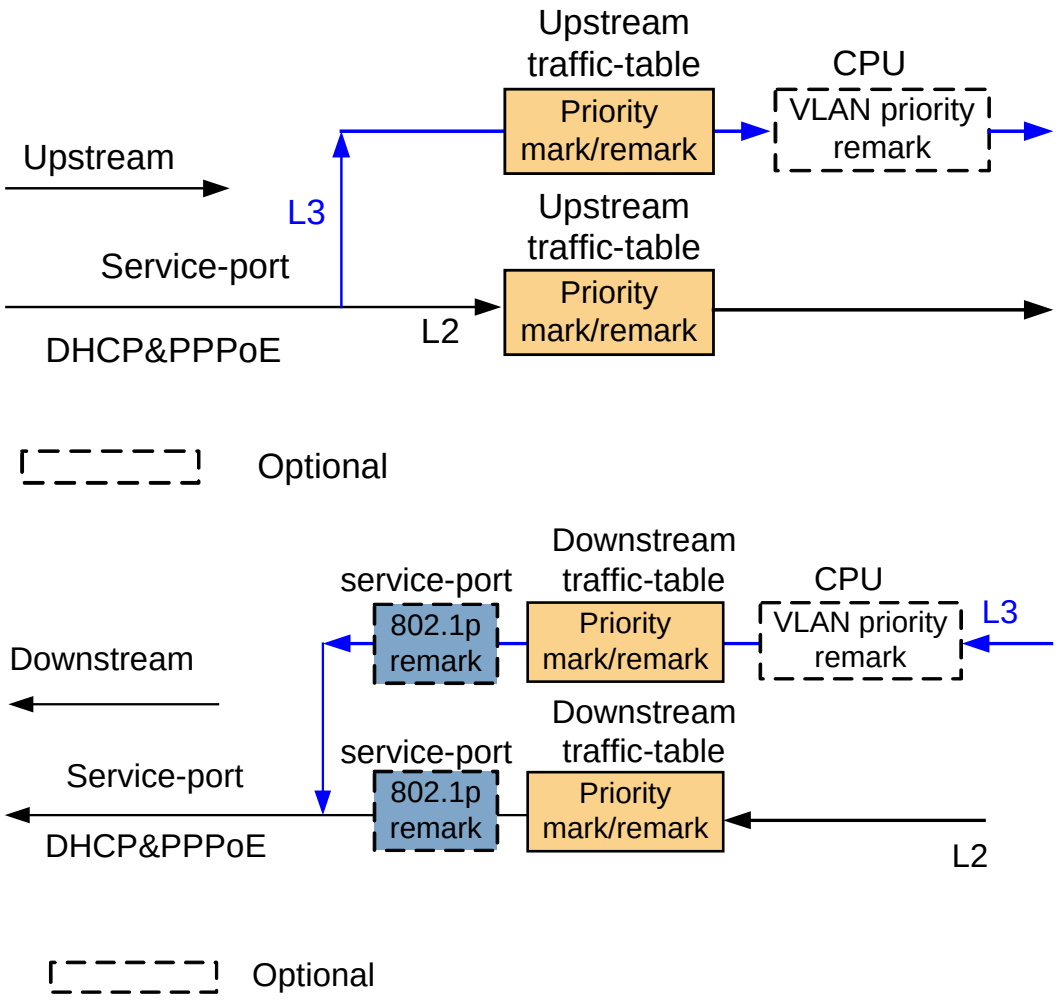
NOTE

Priority processing for PPP session packets is different. For these packets of the service flow created on the GPON/OPGD/OPGE board, if the 802.1p priority is configured for the service flow and the priority processing policy specified in the traffic profile is user-cos, the network-side priority (network-cos) will be copied first.

When DHCP packets work in Layer 3 mode, the priority processing mechanism has the following characteristics:

- Priority processing in the upstream direction is based on the priority processing policy (specify or copy/map priorities) configured in the upstream traffic profile.
- When the **prival** parameter or the **user-cos**/**user-inner-cos**/**user-tos** parameter in the upstream traffic profile is set to 0, the network-side VLAN priority takes effect.
- In the downstream direction, if the 802.1p priority (user-802.1p) is configured for the service flow, remarking is implemented based on the 802.1p priority; if not, priority processing is implemented based on the settings in the downstream traffic profile.
- In the downstream direction, if the network-side priority of the packet is **0**, the packet priority is remarked as the user-side VLAN priority, and then is remarked based on the 802.1p priority of the traffic profile and service flow.
- In the downstream direction, if the priority processing policy specified in the traffic profile is **user-tos**, the user-side priority of the packet is fixedly remarked as **0**.

Figure 14-20 Priority processing for upstream and downstream DHCP & PPPoE packets



Priority Processing for DHCP & PPPoE Packets in the Service-Bundle Scenario

Priority processing for DHCP & PPPoE packets in the service-bundle scenario has the following characteristics:

- The priority of upstream packets is the priority specified for the service-bundle group. The settings in the traffic profile do not apply.
- The priority processing for downstream packets is the same as that in the non-service-bundle scenario.

Priority Processing for DHCPv6 Packets

This topic describes the priority processing for DHCPv6 packets. For detailed processing rules and results, refer to Priority Processing Rules.

Priority Processing for DHCPv6 Packets in the Non-Service-Bundle Scenario

DHCPv6 protocol packets can work in Layer 2 or Layer 3 relay mode. The DHCPv6 protocol packets in this document refer to DHCP Option18/Option37 packets in Layer 2 relay mode and DHCP Layer 3 relay packets.

When DHCPv6 packets work in Layer 2 mode, the priority processing mechanism is the same as that for DHCPv4 packets in Layer 2 mode, which has the following characteristics:

- The service flow priority is irrelevant to the VLAN priority.
- Priority processing in the upstream direction is based on the priority processing policy (specify or copy/map priorities) configured in the upstream traffic profile.
- If the **user-cos**, **user-tos**, or **user-inner-cos** parameter is selected but no such priority exists in the packet (in other words, priority copying fails), the system uses the default 802.1p priority for the packet.
- In the downstream direction, if the 802.1p priority (user-802.1p) is configured for the service flow, remarking is implemented based on the 802.1p priority; if not, priority processing is implemented based on the settings in the downstream traffic profile.

When DHCPv6 packets work in Layer 3 mode, the priority processing mechanism is the same as that for DHCPv4 packets in Layer 3 mode, which has the following characteristics:

- Priority processing in the upstream direction is based on the priority processing policy (specify or copy/map priorities) configured in the upstream traffic profile.
- When the **prival** parameter or the **user-cos**/**user-inner-cos**/**user-tos** parameter in the upstream traffic profile is set to 0, the network-side VLAN priority takes effect.
- In the downstream direction, if the 802.1p priority (user-802.1p) is configured for the service flow, remarking is implemented based on the 802.1p priority; if not, priority processing is implemented based on the settings in the downstream traffic profile.
- In the downstream direction, if the network-side priority of the packet is **0**, the packet priority is remarked as the user-side VLAN priority, and then is remarked based on the 802.1p priority of the traffic profile and service flow.
- In the downstream direction, if the priority processing policy specified in the traffic profile is **user-tos**, the user-side priority of the packet is fixedly remarked as **0**.

Priority Processing for DHCPv6 Packets in the Service-Bundle Scenario

Priority processing for DHCPv6 packets in the service-bundle scenario has the following characteristics:

- The priority of upstream packets is the priority specified for the service-bundle group. The settings in the traffic profile do not apply.
- The priority processing for downstream packets is the same as that in the non-service-bundle scenario.

Priority Processing for IGMP Packets

This topic describes the priority processing for IGMP packets. For detailed processing rules and results, refer to Priority Processing Rules.

The IGMP packets in this document refer to IGMP packets for IPv6 and IPv4. The priority processing applies to IGMP packets for both IPv4 and IPv6 unless otherwise specified.

Priority Processing for IGMP Packets in the Non-Service-Bundle Scenario

The IGMP mode for multicast VLANs can be IGMP snooping or IGMP proxy. The priority processing mechanism for IGMP packets varies with the IGMP mode.

Figure 14-21 describes the priority processing mechanism for IGMP packets in IGMP snooping mode.

- Priority processing in the upstream direction is based on the priority processing policy (specify or copy/map priorities) configured in the upstream traffic profile.
- In the downstream direction, if the 802.1p priority (user-802.1p) is configured for the service flow, remarking is implemented based on the 802.1p priority; if not, priority processing is implemented based on the settings in the downstream traffic profile.

NOTE

The current version does not support the priority-copying policy for IGMP protocol packets. Therefore, when the priority processing policy specified is user-cos or user-tos, the default priority (defaultval) in the downstream traffic profile is used.

Figure 14-21 Priority processing for IGMP packets in IGMP snooping mode

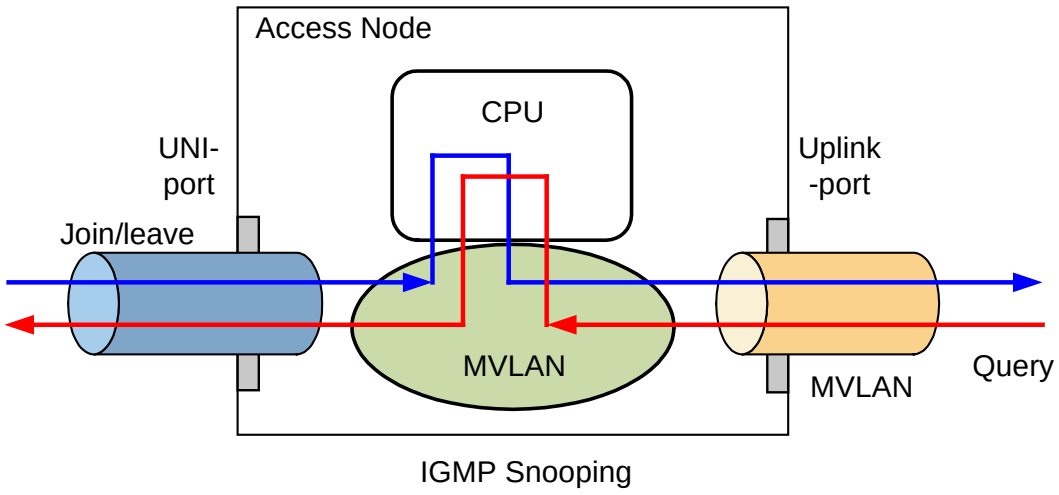


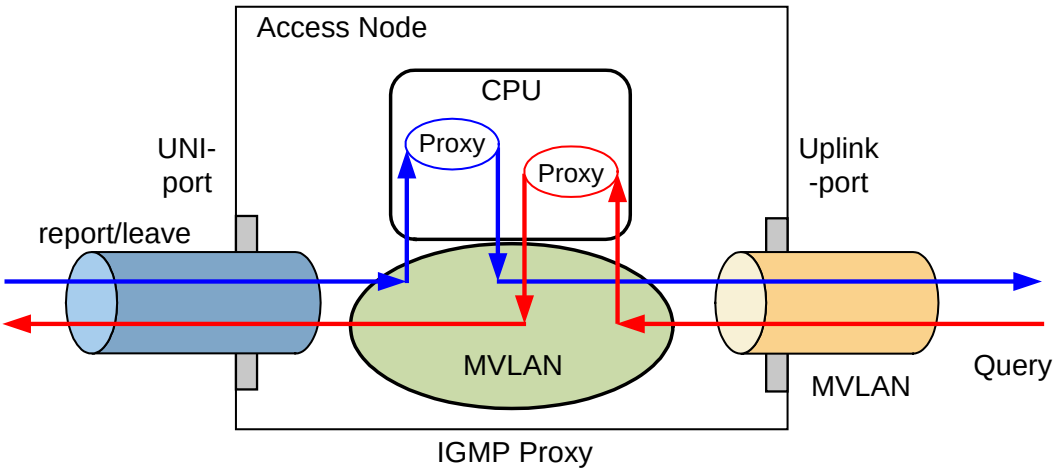
Figure 14-22 describes the priority processing mechanism for IGMP packets in IGMP proxy mode.

- The priority processing in the upstream direction depends on the specified IGMP packet priority. Command for setting the IGMP packet priority: **igmp priority** (IPv4)/**igmp ipv6 priority** (IPv6)
- In the downstream direction, if the 802.1p priority (user-802.1p) is configured for the service flow, remarking is implemented based on the 802.1p priority; if not, priority processing is implemented based on the settings in the downstream traffic profile.

NOTE

The current version does not support the priority-copying policy for IGMP protocol packets. Therefore, when the priority processing policy specified is user-cos or user-tos, the default priority (defaultval) in the downstream traffic profile is used.

Figure 14-22 Priority processing for IGMP packets in IGMP proxy mode



Priority Processing for IGMP Packets in the Service-Bundle Scenario

Priority processing for IGMP packets in the service-bundle scenario has the following characteristics:

- In IGMP proxy mode, the priority processing for upstream and downstream packets is the same as that in the non-service-bundle scenario.
- In IGMP snooping mode, the priority processing mechanism has the following characteristics:
  - The priority of upstream packets is the priority specified for the service-bundle group. The settings in the traffic profile do not apply.
  - The priority processing for downstream packets is the same as that in the non-service-bundle scenario.

Priority Processing for ARP Packets

This topic describes the priority processing for ARP packets. For detailed processing rules and results, refer to Priority Processing Rules.

ARP packets in this document refer to ARP packets sent by access nodes in the upstream and downstream directions. For ARP packets sent by non-access nodes, the priority processing mechanism is the same as that for common data packets.

For ARP packets, the upstream/downstream direction has a meaning that differs from the meaning for other protocol packets:

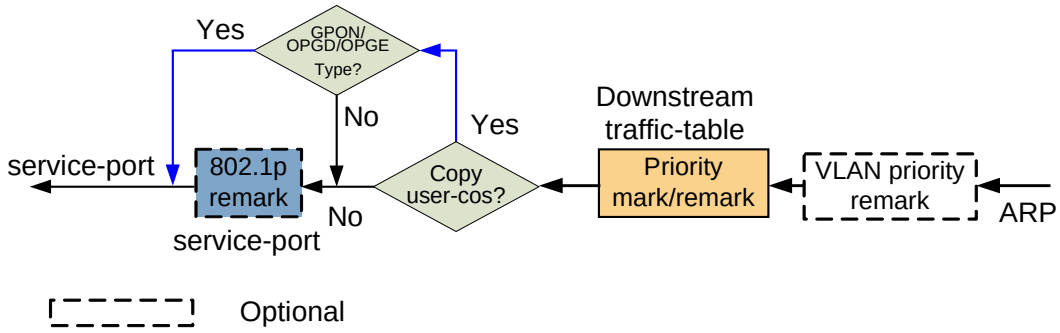
- Upstream direction: In this direction, an access node gives ARP responses or initiates ARP requests to the network side.
- Downstream direction: In this direction, an access node gives ARP responses or initiates ARP requests to the user side.

Priority Processing for ARP Packets in the Non-Service-Bundle Scenario

The priority processing mechanism for ARP packets has the following characteristics:

- In the upstream direction, priority processing depends only on the VLAN priority and is irrelevant to the priority configuration of the service flow.
  - If no VLAN priority is configured, the packet priority will be remarked as 0 (the default priority for ARP packets).
  - If the VLAN priority is configured, the packet priority will be remarked as the VLAN priority.
- Figure 14-23 describes the priority processing in the downstream direction.
  - The priority processing in the downstream direction involves VLAN priority remarking, priority remarking based on the traffic profile, and service flow 802.1p remarking.
  - If the service flow is created on the GPON/OPGD/OPGE board and the priority processing policy specified in the traffic profile is **user-cos**, the 802.1p remarking is skipped. In this case, if the VLAN priority is configured, the packet priority will be remarked based on the VLAN priority; if no VLAN priority is configured, the packet priority will be remarked as 0.
  - If the priority processing policy specified in the traffic profile is user-tos or user-inner-cos, priority copying fails because ARP packets do not have such fields. In this case, the packet priority is remarked as the default value (defaultval) specified in the traffic profile.

Figure 14-23 Priority processing for downstream ARP packets in the non-service-bundle scenario



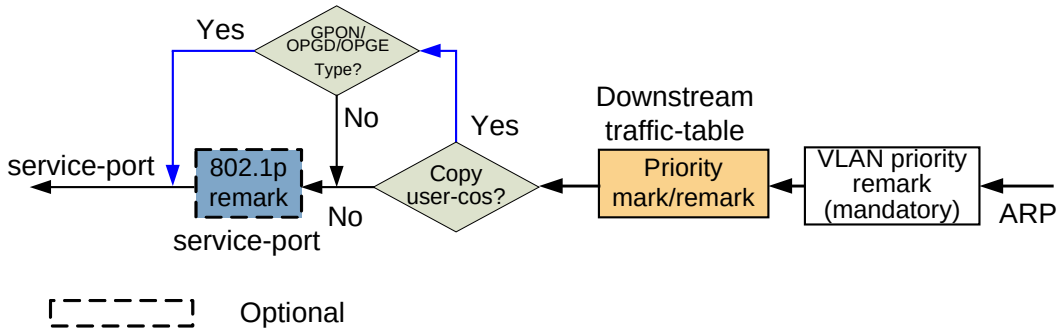
Priority Processing for ARP Packets in the Service-Bundle Scenario

In the upstream direction, priority processing for ARP packets is irrelevant to the priority configuration of the service flow and is the same as that in the non-service-bundle scenario.

In the service-bundle scenario, the priority of ARP packets sent by an access node to the user side must be the bundle-cos of the service flow; otherwise, the ARP packets fail to be received. Therefore, the VLAN priority must be configured.

Figure 14-24 describes the priority processing mechanism for downstream ARP packets in the service-bundle scenario. According to Figure 14-24, the priority processing is almost the same as that in the non-service-bundle scenario. The difference is that the VLAN priority is mandatory in this scenario while it is optional in the non-service-bundle scenario.

Figure 14-24 Priority processing for downstream ARP packets in the service-bundle scenario



Priority Processing for ND Packets

This topic describes the priority processing for neighbor discovery (ND) packets. For detailed processing rules and results, refer to Priority Processing Rules.

ND packets apply to the IPv6 scenario. Similar to the ARP function for IPv4 packets, the ND function is used to determine the link-layer address and reachability of a neighboring node for IPv6 packet forwarding.

Because the function of ND packets is similar to that of ARP packets, the priority processing mechanism for ND packets is also similar to that for ARP packets. For details, refer to Priority Processing for ARP Packets. The difference is that **the default priority for ARP packets is 0 while that for ND packets is 6**.

- **The default priority for ARP packets is 0 while that for ND packets is 6.**
- **The ARP packet does not carry any IP header, and therefore the user-tos copying in the downstream direction is not supported. In this case, the priority of the ARP packet is remarked as the default prority specified in the traffic profile. The ND packet carries an IP header, and therefore the user-tos copying in the downstream direction is supported. In this case, the priority of the ND packet is remarked as 6.**

14.6.5 Configuring the Priority Processing

Context

Priority processing is the basis for equipment or a network to schedule packets. When congestion occurs, equipment or a network schedules packets based on priority.

A priority processing policy is configured in a traffic profile for equipment.

Procedure

Run the **traffic table ip** command to configure priority processing.

```
huawei(config)#traffic table ip index 11
cir<K>|name<K> }:cir
{ cir<U><64,10240000>|off<K> }:off
{ color-policy<K>|priority<K> }:priority
{ prival<U><0,7>|user-cos<K>|user-inner-cos<K>|user-tos<K> }:user-cos
{ defaultval<U><0,7>|mapping-profile<K> }:3
{ inner-priority<K>|priority-policy<K> }:inner-priority
```

```
{ inner-prival<U><0,7>|user-cos<K>|user-inner-cos<K>|user-tos<K> }:user-inner-cos
{ defaultval<U><0,7>|mapping-profile<K> }:3
{ priority-policy<K> }:priority-policy
{ priority-policy<E><Local-Setting,Tag-In-Package,Tag-In-Ingress-Package> }:tag-
in-package

  Command:
        traffic table ip index 11 cir off priority user-cos 3 inner-priority u
ser-inner-cos 3 priority-policy tag-in-package
  Create traffic descriptor record successfully
  -------------------------------------------------
  TD Index           : 11
  TD Name            : ip-traffic-table_11
  Priority           : 3
  Copy Priority       : user-cos
  Mapping Index       : 0
  CTAG Mapping Priority: user-inner-cos
  CTAG Mapping Index   : 0
  CTAG Default Priority: 3
  Priority Policy     : tag-pri
  CIR               : off
  CBS               : off
  PIR               : off
  PBS               : off
  Color policy       : dei
  Referenced Status   : not used
  -------------------------------------------------
```

The parameters for this command are as follows:

- **priority**: specifies the S-VLAN priority policy, including the priority source and queue scheduling policy. The upstream priority determines which queue that upstream packets enter. Valid values for this parameter are as follows:
  - ***prival***: specifies a priority for upstream and downstream packets. When **priority-policy** is set to **Local-Setting**, the priority is determined by the value of **Local-Setting**.
  - **user-cos**: copies the priority from the outer 802.1q tag of an incoming packet as the S-VLAN priority.
  - **user-inner-cos**: copies the priority from the inner 802.1q tag of an incoming packet as the S-VLAN priority.
  - **user-tos**: copies the priority from the IP ToS field of an incoming packet as the S-VLAN priority.
- **inner-priority**: specifies the C-VLAN priority policy. The parameter values are the same as those for the parameter that specifies the S-VLAN priority policy. C-VLAN and S-VLAN can be set independently. Currently, only the SPUA board supports setting of the C-VLAN priority.
- **priority-policy**: specifies the priority policy for queue scheduling. The queue scheduling priority of a packet is generally the same as the priority of the packet. If queue scheduling and packet priorities are different, the queue scheduling priority must be manually specified. Valid values for this parameter are as follows:
  - **Local-Setting**: uses a manually specified priority as the queue scheduling priority.
  - **Tag-In-Package**: uses the priorities of outgoing packets (after VLAN translation) as the queue scheduling priority.

- **Tag-In-Ingress-Package**: uses the priorities of incoming packets (before VLAN translation) as the queue scheduling priority for the downstream direction.

NOTE
For details on this command, see "QoS Configuration" in "QoS Command" of Command Reference.

Step 1 Run the **display traffic table ip** command to verify the configuration result.

----End

Follow-up Procedure

After configuring a traffic profile, run the **service-port** command to bind service flows to the profile to ensure that different priority processing policies are applied to different types of traffic.

14.7 Traffic Policing

Before service providers provide subscribers with specific services, a service level agreement (SLA) is generally assigned, in which all service parameters are defined. To ensure that the user traffic can meet the SLA, the user traffic must be policed.

14.7.1 Introduction

Definition

Traffic policing (also called traffic policy) sets limits on the rate of incoming and outgoing service flow packets. Policing measures packet rates to monitor service flow bursts. Incoming and outgoing packets must meet user-defined conditions and, if they do not meet these conditions, the device that the packets pass through is configured to take countermeasures. For example, if the incoming packet rate exceeds a threshold, the device can take actions such as dropping or coloring the packets (re-setting the packet priorities).

Traffic policing is usually implemented using the committed access rate (CAR). In a PON system, upstream bandwidth conflict between ONUs is resolved by the 2.7.3 DBA technology.

Purpose

Traffic policing enables carriers to achieve the following goals:

- Ensure that user traffic stays within the service level agreement (SLA).
- Ensure service quality by regulating outgoing traffic and suppressing burst traffic.
- Control the rate of broadcast packets using packet suppression.

14.7.2 Basic Concepts

Figure 14-25 shows the basic concepts of traffic policing and Table 14-8 describes these concepts.

Figure 14-25 Key concepts of traffic policing

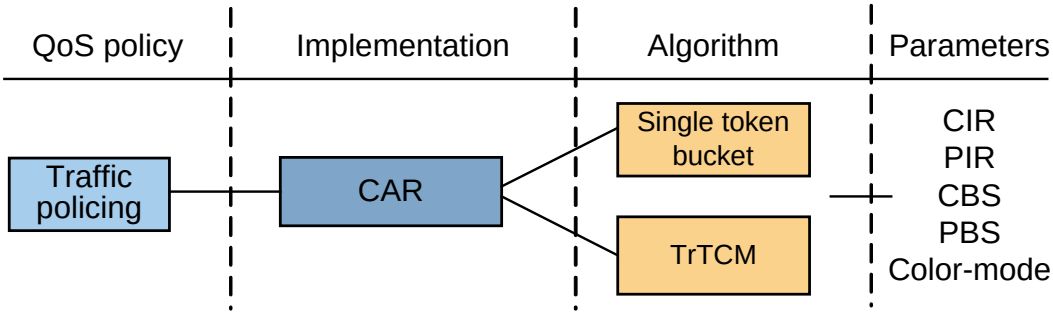


Table 14-8 Description of the basic concepts

Concept	Description
CAR	Stands for committed access rate. It is a technique commonly used to provide the specified rate for specific traffic. And it is widely used to limit the rate of ethernet ports, xDSL ports and xPON ports. CAR is usually implemented using token bucket algorithms.
Token bucket	A token bucket is a container that stores tokens, and it is used for controlling data traffic. A token bucket allows burst data transmission while controlling the traffic. There are two token bucket algorithms: single token bucket algorithm and dual token bucket algorithm. According to the principle of token bucket algorithms, a packet is transmitted when there is equivalent size of tokens in the token bucket. After a packet is transmitted, the number of tokens in the buckets decreases accordingly.
trTCM	Stands for two rate three color marker. It is defined in RFC2698 developed by the Internet Engineering Task Force (IETF). The trTCM algorithm sets the DEI bit for an ethernet packet based on two rates (PIR and CIR) and the burst size to mark the packet green, yellow, or red.
DEI	Stands for drop eligible indicator and is defined by 802.1ad. The DEI field in an ethernet packet consists of 1 bit and is the same as the CFI field defined by 802.1q. This field is used to mark an Ethernet packet with a corresponding color.
CIR	Stands for committed information rate. The unit is bit/s.
PIR	Stands for peak information rate. It provides users the maximum bandwidth when the system is idle. The unit is bit/s.
CBS	Stands for committed burst size. It is used to define the capacity of token bucket C, that is, the maximum burst IP packet size when information is transferred at the committed information rate. The unit is byte.
PBS	Stands for peak burst size. It is used to define the capacity of token bucket P, that is, the maximum burst IP packet size when the information is transferred at the peak information rate. The unit is byte.
Color mode	The trTCM algorithm uses two modes: color-blind and color-aware. In the color-blind mode, the trTCM algorithm assumes that the packet is

Concept	Description
	uncolored. In the color-aware mode, the trTCM algorithm assumes that the packet is pre-colored. The MA5600T/MA5603T/MA5608T support only the color-blind mode.

14.7.3 Implementation Principle: CAR

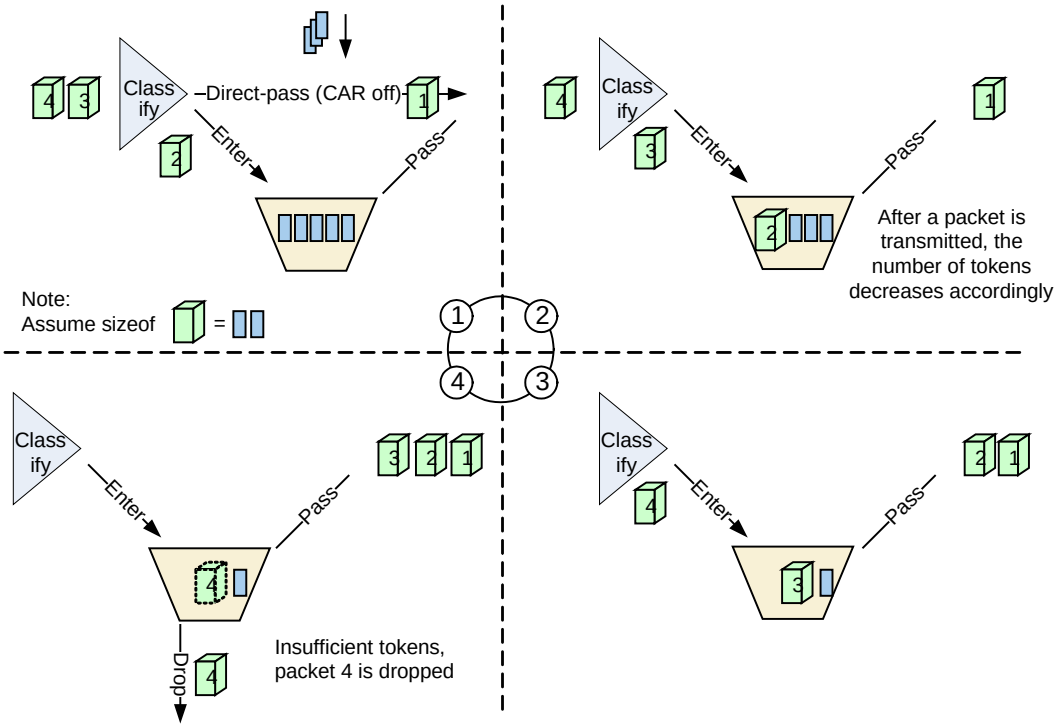
Principle of Single Token Bucket

A packet is transmitted when the number of tokens in the token bucket is sufficient to allow transmission. After a packet is transmitted, the number of tokens decreases accordingly. The details are as follows:

1. Incoming packets are categorized into different traffic classes. If the packets belong to a traffic class for which the rate is limited, for example, packets 2, 3, and 4 in Figure 14-26, the packets are sent to the token bucket for processing. If the packets do not belong to any traffic class, for example, packet 1 in Figure 14-26, the rate of the packets is not limited and the packets are transmitted directly.
2. If the token bucket stores sufficient tokens, packets are transmitted, for example, packets 2 and 3 in Figure 14-26.
3. If the token bucket stores insufficient tokens, packets are dropped, for example, packet 4 in Figure 14-26.
4. The system places tokens in the token bucket at a user-defined rate. When new tokens are generated in the bucket, successive packets can be transmitted.

When the token bucket is full, the system transmits all packets that are the same size as the tokens in the bucket. This process allows for burst transmission. When the token bucket is empty, the system cannot transmit any packet. The system resumes transmissions only after new tokens have been generated. This means that the traffic transmission rate is consistently lower than or equal to the token generation rate, so as to achieve the rate limiting goal.

Figure 14-26 CAR implementation using a single token bucket (process in clockwise direction)



In the figure above, the digits 1, 2, 3, and 4 are the numbers of packets. After a packet passes through the bucket, the number of tokens decreases by the size of the packet. For easy understanding, we suppose that all packets have the same size.

Principle of Dual Token Buckets (trTCM)

The two rate three color marker (trTCM) algorithm defined by RFC2698 is used for traffic policing and marking to achieve effective bandwidth management. If static bandwidth is planned for a network, the trTCM algorithm can at least ensure the basic bandwidth (CIR) for users, and allows users to obtain extra bandwidth (PIR) when the network has sufficient bandwidth. In this way, network resources are used more efficiently.

trTCM principles are as follows:

- trTCM uses two token buckets (P and C buckets). The maximum capacity of the P bucket is represented by the peak burst size (PBS) and that of the C bucket is represented by the committed burst size (CBS). The two token buckets are initially full.

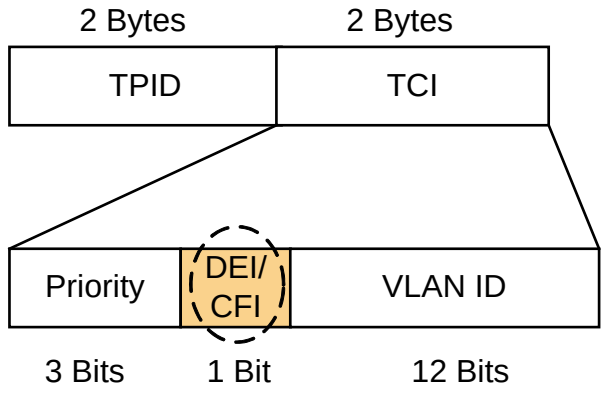
NOTE
PBS must be larger than CBS.

- After a packet passes through the tokens, the number of tokens in the buckets decreases accordingly.
- The system generates tokens in the P bucket at the PIR and in the C bucket at the CIR (PIR >= CIR) per second. The total tokens in the buckets are always less than maximum capacity.
- Based on the size of tokens in the buckets, the system marks the DEI bit (as shown in Figure 14-27) in an incoming packet with a corresponding color (green, yellow, or red). Color coded packets help the system to prevent or manage congestion during data processing.

NOTE

The MA5600T/MA5603T/MA5608T uses the CFI bit defined by 802.1q as the DEI.

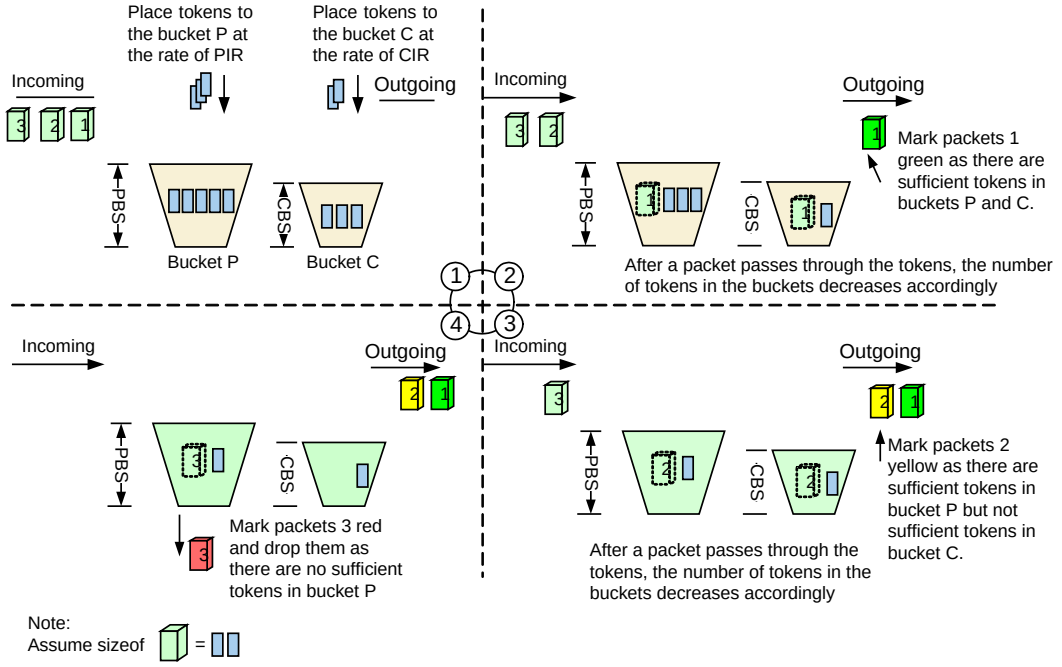
Figure 14-27 DEI bit defined by 802.1ad (same as the CFI bit defined by 802.1q)



As shown in Figure 14-28:

1. If a packet does not exceed CIR, it is marked green (the DEI bit in the packet is set to 0) and is allowed to pass. Packet 1 in the figure is an example.
2. If a packet exceeds CIR but does not exceed PIR, it is marked yellow (the DEI bit in the packet is set to 1) and is allowed to pass. Packet 2 in the figure is an example.
3. If a packet exceeds PIR, the packet is marked red and is directly dropped. Packet 3 in the figure is an example.

Figure 14-28 CAR implementation using the trTCM algorithm (process in clockwise direction)



In the figure above, the digits 1, 2, and 3 are the numbers of packets. After a packet passes through the bucket, the number of tokens decreases by the size of the packet. For easy understanding, we suppose that all packets have the same size.

Enhanced trTCM: Coloring Packets Based on CAR Threshold

The trTCM algorithm ensures that different types of packets are all forwarded when they do not exceed CIR and are forwarded fairly when they reach PIR. The trTCM algorithm polices packets based on the bandwidth and it does not mark colors for packets based on priorities. Therefore, the trTCM algorithm cannot guarantee bandwidth for high-priority services. The enhanced trTCM algorithm can resolve the issue of the trTCM algorithm because it marks colors for packets based on CAR thresholds.

The implementation principle is as follows:

- When the enhanced trTCM algorithm is implemented for equipment, users can set different CAR thresholds for different packets by running the **car-threshold** command, ensuring that a high CAR threshold is set for high-priority packets.

  NOTE
  The CAR threshold is the percentage by which tokens in the C and P buckets have decreased. The percentage of the remaining tokens in the buckets is obtained by subtracting the CAR threshold from 1.

- When packets of different priorities sequentially pass the P bucket and C bucket:
  - If a packet of a certain priority in the P bucket exceeds the CAR threshold, the packet is directly dropped. Otherwise, the packet enters the C bucket.
  - If a packet of a certain priority in the C bucket exceeds the CAR threshold, the packet is marked yellow; otherwise, it is marked green.
- After a packet passes through the tokens, the number of tokens in the buckets decreases accordingly.

As shown above, enhanced trTCM ensures that packets with higher priorities are marked green and their bandwidths are guaranteed.

As shown in Figure 14-29, there are two token buckets C and P with depths CBS and PBS respectively. The two buckets have token counts Tc and Tp, respectively. Tp(t) and Tc(t) represent the number of tokens in the P and C buckets respectively at time t. Initially (t = 0), the P and C buckets are full, that is, Tp(0) = PBS and Tc(0) = CBS. Then, Tp increases by one for PIR times per second until reaching PBS and Tc increases by one for CIR times per second until reaching CBS. Lvl(i) represents the percentage of the remaining tokens corresponding to a specific priority.

When an i-priority packet with length of B bytes arrives at time t:

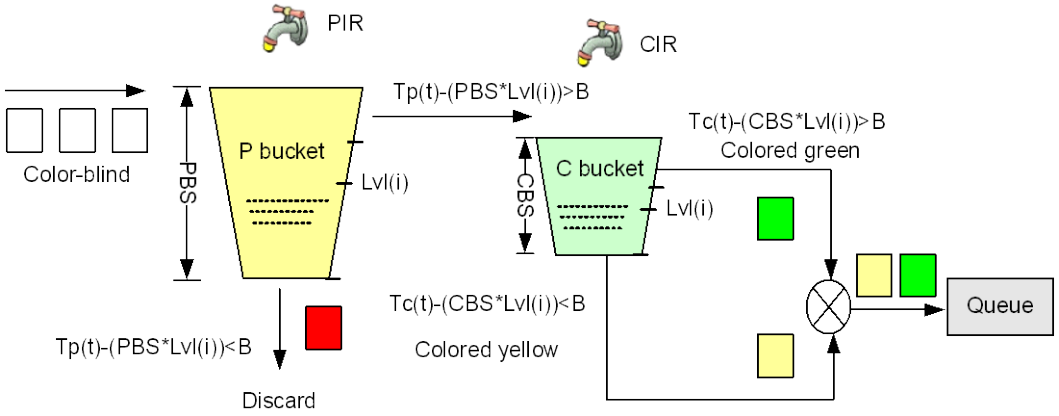
1. If Tp(t) - (PBS*Lvl(i)) < B, the device drops the packet; otherwise, the device sends the packet to the C bucket, and Tp(t) decreases by B bytes.
2. If Tc(t) - (CBS*Lvl(i)) < B, the device marks the packet yellow and Tc(t) decreases by B bytes; otherwise, the device marks the packet green and Tc(t) decreases by B bytes.

For example, when the CAR threshold is 12% for 0-priority packets and is 100% for 6-priority packets, the enhanced trTCM algorithm enables the device to behave as follows:

- When the incoming packets enter the P bucket at a certain time, if the percentage of the remaining tokens in the bucket is less than 88%, the device drops 0-priority packets and sends 6-priority packets to the C bucket. When the incoming packets enter the C bucket at a certain time, if the percentage of the remaining tokens in the bucket is less than 88%, the device marks 0-priority packets yellow and 6-priority packets green.
- When the incoming packets enter the P bucket at a certain time, if there is no token in the bucket, the device also drops 6-priority packets. When the incoming packets enter the C

bucket at a certain time, if there is no token in the bucket, the device drops 0-priority packets and marks 6-priority packets yellow.

Figure 14-29 Coloring packets based on CAR threshold



14.7.4 Traffic Policing Mode

As shown in Table 14-9, the MA5600T/MA5603T/MA5608T support multiple traffic policies:

- Rate limiting based on service flows
- Rate limiting based on port+priority
- Rate limiting based on GEM port+CoS group
- Rate limiting based on service flow+priority
- Rate limiting based on groups of service flows
- Rate limiting based on port+VLAN
- Rate limiting based on ONT
- Rate limiting based on Ethernet port on ONT

Table 14-9 Traffic policing mode and token bucket algorithm

Traffic Policy	Description	Token Bucket Algorithm
Rate limiting based on service flows	When configuring service flows, you can limit the rate of the service flows by binding upstream and downstream service flows to an upstream and downstream profile, respectively. If there is no limit on the rate, you can configure the upstream and downstream profiles as **car-off**. Rate limiting based on service flows and rate limiting based on port+CoS cannot be applied to the same board or port at the same time.	trTCM

Traffic Policy	Description	Token Bucket Algorithm
	The configuration commands are: • **car-mode service-port** • **service-port** • **traffic table ip**	
Rate limiting based on port+priority	Rate limiting based on port+CoS is used for xDSL applications. Rate limiting based on GEM port+CoS is used for GPON applications. Rate limiting based on service flows and rate limiting based on port+CoS cannot be applied to the same board or port at the same time. The configuration commands are: • **car-mode port-cos** • **car-port *portid* cos** • **car-port** *portid* **ont** *ontid* **gemindex** *gemindex* **cos**	trTCM
Rate limiting based on GEM port+CoS group	One GEM port can carry services with different priorities. After the priority of the specific service is added to the CoS group, rate limiting is performed based on this CoS group. This meets the carrier requirement that rates of various services carried on the same GEM port are limited. The configuration commands are: • **car-mode port-cos** • **cos-group-table** • **car-port** *portid* **ont** *ontid* **gemindex** *gemindex* **cos-group-table** *table-index* • **car-port** *portid* **ont** *ontid* **gemindex** *gemindex* **cos-group** *group-id*	trTCM
Rate limiting based on service flow+priority	When configuring service flows, you can limit the rate of the service flows by binding upstream and downstream service flows to an upstream and downstream profile, respectively. If there is no limit on the rate, you can configure the upstream and downstream profiles as **car-off** (system default). The configuration command is: **car-threshold**.	Enhanced trTCM
Rate limiting based on groups of service flows	You can bind multiple upstream or downstream service flows into one group and apply the same CAR to the group.	Single token bucket Note: This rate limiting policy applies when the

Traffic Policy	Description	Token Bucket Algorithm
	The configuration command is: **car-group**.	single token bucket algorithm and the rate limiting policy is bound to the TrTCM traffic profile. In the TrTCM traffic profile, CIR must be equal to PIR; otherwise, the preconfigured PIR will supersede the specified rate limiting policy.
Rate limiting based on port+VLAN	You can limit the rates for packets with specified VLANs on specified ports. Only the SPUA board supports this traffic policy. The configuration command is: **car-port** *portid* **vlan**.	Single token bucket Note: This rate limiting policy applies when the single token bucket algorithm and the rate limiting policy is bound to the TrTCM traffic profile. In the TrTCM traffic profile, CIR must be equal to PIR; otherwise, the preconfigured PIR will supersede the specified rate limiting policy.
Rate limiting based on ONT	You can limit the rates for downstream packets of a specified ONU. The configuration command is: **traffic-limit ont**	Single token bucket Note: This rate limiting policy applies when the single token bucket algorithm and the rate limiting policy is bound to the TrTCM traffic profile. In the TrTCM traffic profile, CIR must be equal to PIR; otherwise, the preconfigured PIR will supersede the specified rate limiting policy.
Rate limiting based on Ethernet port on ONT	You can limit the rates for upstream and downstream packets of specified Ethernet ports on an ONT. The configuration command is: **ont port car**	Single token bucket Note: This rate limiting policy applies when the single-token bucket algorithm and the rate limiting policy is bound to the TrTCM traffic profile. In the TrTCM traffic profile, CIR must be equal to PIR;

Traffic Policy	Description	Token Bucket Algorithm
		otherwise, the preconfigured CIR will supersede the specified rate limiting policy.
Traffic suppress	Using this policy, the system suppresses broadcast, unknown multicast, and unknown unicast packets for inbound ports. The traffic suppression prevents these packets from consuming excessive network resources and therefore protects the network from congestion. The configuration command is: **traffic-suppress**	-

14.7.5 Configuring the Traffic Policing

Configuring Rate Limitation Based on Service Port

This topic describes how to limit rate on a specific service flow (on behalf of a different type of service or users) through the traffic parameters defined in the IP traffic profile.

Context

- The system has seven default IP traffic profiles with the IDs of 0-6. You can run the **display traffic table** command to query the traffic parameters of the default traffic profiles.
- It is recommended that you use the default traffic profiles. A new IP traffic profile is created only when the default traffic profiles cannot meet the requirements.

Procedure

Configure the rate limitation mode to be service-port.

In the service board mode, run the **car-mode** [*portlist*] **service-port** command to set the rate limitation mode of board or port to be service-port.

Step 1 Configure the parameters of IP traffic profile.

When the default traffic profiles cannot meet the requirements, run the **traffic table ip** command to configure the IP traffic profile.

Table 14-10 shows the traffic parameters related to rate limitation in IP traffic profile.

Table 14-10 Traffic parameters and parameter Description

Item	Parameter Description
cir *cir*	Stands for committed information rate. It ensures that users are assigned bandwidth even when the system is busy. The unit is bit/s.

Item	Parameter Description
	CIR is mandatory, and must be an integer multiple of 64. If it is not, the value is rounded down to a nearest integer multiple of 64 but cannot be smaller than 64.
cbs *cbs*	Stands for committed burst size. It indicates the maximum capacity that a token bucket buffers tokens. The unit is byte. CBS is optional. If the parameter is not specified, it can be obtained by the formula min(2000 + 32*cir,10240000).
pir *pir*	Stands for peak information rate. It provides users the maximum bandwidth when the system is idle. The unit is bit/s. PIR is optional. If the parameter is not specified, it can be obtained by the formula min(2*cir, 10240000). The pir cannot be smaller than cir.
pbs *pbs*	Stands for peak burst size. It indicates the maximum capacity that a token bucket buffers tokens. The unit is byte. PBS is optional. If the parameter is not specified, it can be obtained by the formula min(2000+32*pir, 10240000).

NOTE

The system marks the service packets with colors according to the CIR and PIR parameters. To be specific, for the packets whose rate is equal to or lower than CIR, the system marks them as green (allowed to pass). For the packets whose rate is higher than CIR and lower than PIR, the system marks them as yellow (allowed to pass). For the packets whose rate is higher than PIR, the system marks them as red, and drops such packets.

Step 2 Run the **service port** command to bind the specified IP traffic profile.

----End

Example

Assume that the CIR is 2048 kbit/s, 802.1p priority of the outbound packet is 6, and the scheduling policy of the inbound packet is Tag-In-Package. To add traffic profile 9 with these settings, do as follows:

```
huawei(config)#traffic table ip index 9 cir 2048 priority 6 priority-policy
tag-In-Package
  Create traffic descriptor record successfully
  ------------------------------------------------
  TD Index           : 9
  TD Name            : ip-traffic-table_9
  Priority           : 6
  Copy Priority       : -
  Mapping Index       : -
  CTAG Mapping Priority: -
  CTAG Mapping Index   : -
  CTAG Default Priority: 0
  Priority Policy     : tag-pri
  CIR                : 2048 kbps
  CBS                : 67536 bytes
  PIR                : 4096 kbps
```

```
  PBS               : 133072 bytes
  Color policy       : dei
  Referenced Status   : not used
  Referenced Status   : not used
  ---------------------------------------------
huawei(config)#display traffic table ip index 9
  ---------------------------------------------
  TD Index          : 9
  TD Name           : ip-traffic-table_9
  Priority          : 6
  Copy Priority      : -
  Mapping Index      : -
  CTAG Mapping Priority: -
  CTAG Mapping Index  : -
  CTAG Default Priority: 0
  Priority Policy     : tag-pri
  CIR               : 2048 kbps
  CBS               : 67536 bytes
  PIR               : 4096 kbps
  PBS               : 133072 bytes
  Color policy       : dei
  Referenced Status   : not used
  ---------------------------------------------
```

Configuring Rate Limitation Based on Port+CoS

This topic describes how to configure Rate limiting based on port+CoS so that different IP traffic profiles can be specified for the traffic streams that have different 802.1p priorities on a port.

Prerequisites

The IP traffic profile is configured. For details, see Configuring Rate Limitation Based on Service Port.

Context

- Rate limiting based on service ports conflicts with Rate limiting based on port+CoS. By default, the system supports Rate limiting based on service ports.
- If service ports are configured on the board, the rate limiting mode of the board cannot be changed.

Procedure

Configure the rate limitation mode to be port-cos.

In the service board mode, run the **car-mode port-cos** command to configure the rate limiting mode of the service board to rate limiting based on port+CoS.

Step 1 Configure rate limitation for service flows with specified 802.1p priority.

Run the **car-port** command to specify the 802.1p priority for the port, and bind an IP traffic profile to the traffic streams that meet the specified 802.1p priority.

When rate limiting based on port+CoS is selected for a board, pay attention to the following points:

- For a non-xPON board, you can bind the corresponding traffic profile in the inbound/outbound direction according to a CoS value of a port on the board.
- For a GPON board, you can bind the corresponding traffic profile in the inbound/outbound direction according to a CoS value of a GEM port on the board.

----End

Example

To configure GEM port 130 on port 0 of the GPON board in slot 0/2, and bind traffic profile 2 to the packets with priority 7, do as follows:

```
huawei(config)#interface gpon 0/2
huawei(config-if-gpon-0/2)#car-mode port-cos
huawei(config-if-gpon-0/2)#car-port 0 gemport 130 cos 7 inbound 2 outbound 2
huawei(config-if-gpon-0/2)#display car-mode
  The CAR mode of the board: port-cos
huawei(config-if-gpon-0/2)#display car-port 0 gemport 130
  ---------------------------------------------
  Port GEM port CoS Inbound-index Outbound-index
  ---------------------------------------------
    0     130   7           2            2
  ---------------------------------------------
```

To configure port 0 of theVDSL2 board in slot 0/2, and bind traffic profile 3 to the packets with priority 3, do as follows:

```
huawei(config)#interface vdsl 0/2
huawei(config-if-vdsl-0/2)#car-mode port-cos
huawei(config-if-vdsl-0/2)#car-port 0 cos 3 inbound 3 outbound 3
huawei(config-if-vdsl-0/2)#display car-mode
  The CAR mode of the board: port-cos
huawei(config-if-vdsl-0/2)#display car-port 0
  ---------------------------------------------
  Port  CoS Inbound-index Outbound-index
  ---------------------------------------------
   0    3           3            3
  ---------------------------------------------
```

Configuring User-based Rate Limitation

In the user-based rate limitation, the VoIP, IPTV service, and Internet access service of each user share a total user bandwidth. When there is no voice or IPTV service, the Internet access service can hold a burst of the total user bandwidth so that the total user bandwidth can be managed in a unified manner.

Context

When the user uses the Triple play service, the VoIP, IPTV service, and Internet access service of each user share a total user bandwidth. All services of the user hold the total user bandwidth, and the service with the highest CoS priority is ensured first. When other services

carry no traffic, each service can hold a burst of the total user bandwidth. The multicast bandwidth is determined by the bandwidth of demanded programs. The total bandwidth of demanded programs cannot exceed the total user bandwidth.

NOTE
PON multicast services do no support rate limitation in CAR-group mode.

Procedure

- For PON access users.
  - In the user-based rate limitation, multiple service ports of a user are added to a rate-limited group. Through the QoS strategy applied on the rate-limited group, the total user bandwidth is ensured on the basis that the committed information rate (CIR) and peak information rate (PIR) of each service are ensured, and each service is allowed to hold a burst of the total user bandwidth.

  a. Run the **traffic table ip** command to create an IP traffic profile to configure the CoS priority of each service and ensure the CIR and PIR.
     - The CoS priorities of services are VoIP, IPTV service, and Internet access service in a descending order.
     - In the IP traffic profile used by the rate-limited group, the PIR must be equal to or larger than the sum of CIRs of all services in other IP traffic profiles.

  b. Run the **service-port** command to create service ports of the VoIP, IPTV service, and Internet access service, using the IP traffic profile created in a.

  c. Run the **car-group** command to create the rate-limited group of service ports to manage the total user bandwidth of multiple services.
     - To ensure the user bandwidth, the PIR of the rate-limited group must be equal to or larger than the sum of CIRs of all services in the rate-limited group.
     - The PIR is equal to the total user bandwidth. In the case that any two services carry no traffic, the third service can hold a burst of the total user bandwidth.

  d. Run the **car-group add-member service-port** command to add service ports to the rate-limited group.

     Pay attention to the following points when adding service ports to the rate-limited group:
     - Only service ports of the same PON port can be added to the same rate-limited group.
     - For Type C and Type D, only service ports of the same ONT can be added to the same rate-limited group.
     - One service port cannot be added to multiple rate-limited groups.
     - A maximum of eight service ports can be added to a rate-limited group.

- For ADSL2+ and VDSL access users.

  Each port corresponds to a user. By limiting the upstream/downstream rate of the port, set the maximum upstream/downstream rate to the total user bandwidth. All services of the user hold the total user bandwidth, and the service with the highest CoS priority is ensured first. When other services carry no traffic, each service can hold a burst of the total user bandwidth.

  a. Set the maximum upstream/downstream rate to the total user bandwidth.
     - For the ADSL2+ access mode:

1) Run the **adsl line-profile quickadd** command to quickly add an ADSL2+ line profile, or run the interactive **adsl line-profile add** command to add an ADSL2+ line profile.
2) Run the **adsl channel-profile quickadd** command to quickly add an ADSL2+ channel profile, or run the interactive **adsl channel-profile add** command to add an ADSL2+ channel profile. In the channel profile, configure the maximum upstream and downstream rates to limit the user bandwidth.
3) Run the **adsl line-template quickadd** command to quickly add an ADSL+ line template, or run the interactive **adsl line-template add** command to add an ADSL2+ line template.

- For the VDSL (common mode) access mode:
  1) Run the **vdsl line-profile quickadd** command to quickly add a VDSL line profile, or run the interactive **vdsl line-profile add** command to add a VDSL line profile.
  2) Run the **vdsl channel-profile quickadd** command to quickly add a VDSL channel profile, or run the interactive **vdsl channel-profile add** command to add a VDSL channel profile. In the channel profile, configure the maximum upstream and downstream rates to limit the user bandwidth.
  3) Run the **vdsl line-template quickadd** command to quickly add a VDSL2 line template, or run the interactive **vdsl line-template add** command to add a VDSL2 line template.
- For the VDSL (TI mode) access mode:
  1) Run the **vdsl service-profile quickadd** command to quickly add a VDSL2 service profile, or run the interactive **vdsl service-profile add** command to add a VDSL2 service profile.
  2) Run the **vdsl spectrum-profile quickadd** command to quickly add a VDSL2 spectrum profile, or run the interactive **vdsl spectrum-profile add** command to add a VDSL2 spectrum profile.

b. Run the **traffic table ip** command to create an IP traffic profile to configure the CoS priority of each service and ensure the CIR and PIR. The PIR is equal to the total user bandwidth. When other services carry no traffic, each service can hold a burst of the total user bandwidth.

   The CoS priorities of services are VoIP, IPTV service, and Internet access service in a descending order.

c. Run the **service-port** command to create service ports of the services, using the IP traffic profile created in b.

d. Run the **queue-scheduler strict-priority** command to configure queue scheduling mode of the port to strict priority queue scheduling.

----End

Example

Assume that under GPON port 0/2/1, the user with the ONT 1 is provided with the VoIP, IPTV, and Internet access services. Set the total user bandwidth to 10 Mbit/s, add rate-limited group 0, add service ports 100, 101, and 102 of the user to rate-limited group 0, and use traffic profile 30 to control traffic of rate-limited group 0. In the case that any two services carry no traffic, the third service can hold a burst of the total user bandwidth. To perform such a configuration with the following parameters, do as follows:

- Service port 100 of the Internet access service uses traffic profile 10, with the CIR 2 Mbit/s and the 802.1p priority 4.
- Service port 101 of the VoIP service uses traffic profile 11, with the CIR 1 Mbit/s and the 802.1p priority 6.
- Service port 102 of the IPTV service uses traffic profile 12, with the packet rate not limited and the 802.1p priority 5.

```
huawei(config)#traffic table ip index 10 cir 2048 pir 10240 priority 4 priority-policy
local-Setting
huawei(config)#service-port 100 vlan 2 gpon 0/2/1 ont 1 gemport 4 multi-service user-vlan
20 rx-cttr 10 tx-cttr 10
huawei(config)#traffic table ip index 11 cir 1024 pir 10240 priority 6 priority-policy
local-Setting
huawei(config)#service-port 101 vlan 2 gpon 0/2/1 ont 1 gemport 5 multi-service user-vlan
30 rx-cttr 11 tx-cttr 11
huawei(config)#traffic table ip index 12 cir off priority 5 priority-policy
local-Setting
huawei(config)#service-port 102 vlan 2 gpon 0/2/1 ont 1 gemport 6 multi-service user-vlan
40 rx-cttr 12 tx-cttr 12
huawei(config)#traffic table ip index 30 cir 10240 pir 10240 priority 3 priority-policy
local-Setting
huawei(config)#car-group 0 inbound traffic-table index 30 outbound traffic-table index
huawei(config)#car-group 0 add-member service-port 100-102
huawei(config)#display car-group 0

Command:
        display car-group 0
 ------------------------------------------------------------------------------
                                                 Inbound   Outbound
 GroupID                                Member List    Index    Index
 ------------------------------------------------------------------------------
      0                               100,101,102      10       10
 ------------------------------------------------------------------------------
Total: 1
```

Assume that under ADSL port 0/3/1, a user is provided with the VoIP, IPTV, and Internet access services. Set the total user bandwidth to 10 Mbit/s. In the case that any two services carry no traffic, the third service can hold a burst of the total user bandwidth. To perform such a configuration with the following parameters, do as follows:

- Service port 100 of the Internet access service uses traffic profile 10, with the CIR 2 Mbit/s and the 802.1p priority 4.
- Service port 101 of the VoIP service uses traffic profile 11, with the CIR 1 Mbit/s and the 802.1p priority 6.
- Service port 100 of the IPTV service uses traffic profile 12, with the packet rate not limited and the 802.1p priority 5.

```
huawei(config)#adsl line-profile quickadd 10
huawei(config)#adsl channel-profile quickadd 10 rate 32 32 10240 32 32 6000
huawei(config)#adsl line-template quickadd 10 channel1 10 10 60 channel2 10
huawei(config)#interface adsl 0/3
huawei(config-if-adsl-0/3)#deactivate 1
huawei(config-if-adsl-0/3)#activate 1 template-index 10
```

```
huawei(config-if-adsl-0/3)#quit
huawei(config)#traffic table ip index 10 cir 2048 pir 10240 priority 4 priority-policy local-Setting
huawei(config)#service-port 100 vlan 2 adsl 0/3/1 vpi 0 vci 35 multi-service user-vlan 20 rx-cttr 10 tx-cttr 10
huawei(config)#traffic table ip index 11 cir 1024 pir 10240 priority 6 priority-policy local-Setting
huawei(config)#service-port 101 vlan 2 adsl 0/3/1 vpi 0 vci 35 multi-service user-vlan 30 rx-cttr 11 tx-cttr 11
huawei(config)#traffic table ip index 12 cir off priority 5 priority-policy local-Setting
huawei(config)#service-port 102 vlan 2 adsl 0/3/1 vpi 0 vci 35 multi-service user-vlan 40 rx-cttr 12 tx-cttr 12
huawei(config)#queue-scheduler strict-priority
```

Configuring Rate Limitation Based on Port+VLAN

After configuring rate limitation based on port+VLAN, you can specify different IP traffic profiles for different VLAN packets carried on the same port.

Prerequisites

- A proper IP traffic profile must be created and the index of the IP traffic profile to be used must be confirmed. For details about the configuration method, see Configuring Rate Limitation Based on Service Port.
- Currently, only the SPUA board supports rate limitation based on port+VLAN.

Procedure

In the global config mode, run the **interface eth** command to enter the ETH mode.

Step 1 Run the **car-port *portid* vlan** command to configure rate limitation based on port+VLAN.

This command can be used to configure IP traffic profiles for the packets in the specified VLAN range on the specified port, implementing inbound and outbound rate limitation.

----End

Example

To configure port 0 on the SPUA board in slot 0/2, and use traffic profile 6 for controlling the packets with VLAN 10, do as follows:

```
huawei(config)#display traffic table ip index 6
  ------------------------------------------------
  TD Index           : 6
  TD Name            : ip-traffic-table_6
  Priority           : 6
  Copy Priority       : user-cos
  Mapping Index       : 0
  CTAG Mapping Priority: -
  CTAG Mapping Index   : -
  CTAG Default Priority: 0
  Priority Policy     : tag-pri
```

```
  CIR               : off
  CBS               : off
  PIR               : off
  PBS               : off
  Color policy       : dei
  Referenced Status   : used
  ------------------------------------------------
huawei(config)#interface eth 0/2
huawei(config-if-eth-0/2)#car-port 0 vlan 10  inbound 6 outbound 6
```

Configuring GPON Rate Limitation

This topic describes how to configure rate limitation for GPON services, thereby providing differentiated quality of service (QoS) for various GPON services.

Context

- There are multiple methods of rate-limiting GPON services, for example, rate-limiting downstream traffic by using an IP traffic profile and ACL rules, rate-limiting the ONT upstream bandwidth by using a DBA profile, and rate-limiting the GEM port and GEM port traffic on an ONT.
- Rate limitation on GPON services can be performed on the OLT and the ONT concurrently. If more than one rate limitation modes are configured in the system, the minimum rate prevails.
- Which method of rate-limiting the ONT upstream bandwidth is used depends on the ONT capability. Specifically, if an ONT supports various rate limitation methods and the ONT upstream traffic is small (for example, FTTH service), a DBA profile is a best choice to rate-limit the ONT upstream traffic. If a T-CONT carries upstream traffic for multiple users (for example, FTTB/FTTC service), rate limitation on GEM port is generally used to prevent a user from occupying bandwidth for a long time. If the priority of user packets is trustable (for example, an enterprise user), priority queue (PQ) scheduling is generally used.

Procedure

- Perform rate limitation on the OLT.
  - Rate limitation using an IP traffic profile includes two modes. For details, see Configuring Rate Limitation Based on Service Port, and Configuring Rate Limitation Based on Port+CoS.
  - Performing rate limitation by configuring an ACL rule can control the traffic matching the ACL rule. For details, see Controlling the Traffic Matching an ACL Rule.
- Perform rate limitation on the ONT.

NOTE

- In the case of an MxU device, rate limitation can be performed on downstream traffic of a service port or a port by configuring an IP traffic profile. For details, see MxU manuals.
- In the case of H805GPBD board, you can run the **traffic-limit ont** command to limit the traffic of downstream packets on a specified ONT. The system limits the traffic of downstream packets on an ONT by using the shaping function and buffers the packets that exceed the limit (that is the PIR parameter in traffic profile) and transmits them at a proper time (such as during periodic checks). This reduces packet drop and at the same time complies with traffic features.

a. Run the **dba-profile add** command to add a DBA profile. The DBA profile is used to schedule the ONT upstream bandwidth properly, achieving the best bandwidth utilization.

   A DBA profile supports five types (Type1 to Type5). Generally, Services with a higher priority adopts Type1 or Type2 DBA profiles and services with a lower priority adopts Type3 or Type4 DBA profiles. Table 14-11 shows the features of the DBA profile of each type.

Table 14-11 The features of the DBA profile

Profile Type	Features
Type1	Indicates the fixed bandwidth. After the DBA profile of Type1 is bound, the system assigns a specified bandwidth, regardless of whether there is upstream traffic.
Type2	Indicates the assured bandwidth. After the DBA profile of Type2 is bound, the system meets the bandwidth requirements if the upstream traffic does not exceed a specified value. When there is no upstream traffic, the system does not assign any bandwidth.
Type3	Indicates the hybrid of assured bandwidth and non-assured bandwidth. The DBA profile of Type3 specifies an assured value and non-assured value. After assigning the fixed bandwidth and assured bandwidth, the system assigns the remaining bandwidth (if any) to the user bound with the DBA profile of Type3 (the assigned bandwidth does not exceed the non-assured bandwidth).
Type4	Indicates the best-effort bandwidth. The DBA profile of Type4 just specifies a maximum value. After the DBA profile of Type4 is bound, its priority for obtaining the bandwidth is the lowest. That is, after assigning the fixed bandwidth, assured bandwidth, and non-assured bandwidth, the system assigns the remaining bandwidth (if any) to the user bound with the DBA profile of Type4 (the assigned bandwidth does not exceed the maximum value).
Type5	Indicates the hybrid bandwidth. The preceding four types of values need to be specified.

b. Run the **ont-lineprofile gpon** command to add a GPON ONT line profile, and then enter the GPON ONT line profile mode.

c. Run the **tcont** command to bind a T-CONT to the DBA profile.

   It is recommended that one service type use one T-CONT and different T-CONTs be planned with different bandwidth assurance types.

d. Run the **qos-mode** command to configure a QoS mode of the GPON ONT line profile to ensure that the QoS mode is the same as that of the GEM port.

   By default, the QoS mode of the GPON ONT line profile (that is, the ONT scheduling mode) is priority queue (PQ). The QoS mode includes:

   - gem-car: Indicates the rate limitation mode based on the GEM port of the T-CONT. Rate limitation is performed on a specified GEM port in the ONT upstream direction. To select the gem-car mode, set **gem add** to **gem-car**. The maximum traffic is determined by the DBA profile bound to the GEM port. If a T-CONT contains multiple GEM ports, the scheduling mechanism of packets

between multiple GEM ports depends on the default scheduling mechanism of the ONT.

- flow-car: Indicates the rate limitation mode based on traffic streams of a GEM port. Rate limitation is performed on a specified traffic stream in the ONT upstream direction. To select the flow-car mode, set **gem mapping** to **flow-car**. The maximum traffic is determined by the DBA profile bound to the traffic stream. Flow-car is more specific than gem-car. After rate limitation based on traffic streams is performed, traffic is scheduled in the T-CONT queue. The scheduling mechanism depends on the default scheduling mechanism of the ONT. Before configuring flow-car, make sure that the required traffic profile is created by running the **traffic table ip** command.

NOTE

The traffic stream in this topic refers to the service channel between an ONT and OLT. It is different the service port created by running the **service-port** command.

- priority-queue: Indicates the PQ mode based on the GEM port of the T-CONT. Traffic is scheduled based on PQ between multiple GEM ports in the ONT upstream direction. To select priority-queue mode, set **gem add** to **priority-queue**. By default, the system supports eight (0–7) queues. Queue 7 has the highest priority and services of queue 7 are preferentially guaranteed. The maximum traffic is determined by the DBA profile to which the T-CONT is bound.

e. Run the **commit** command to make the profile configuration take effect. The configuration of the line profile takes effect only after you run this command.

----End

Example

Assume that:

- A user under ONT 1 connected to GPON port0/2/1 requires 2 Mbit/s high-speed Internet access service.
- The priority of user packets is trustable. The PQ scheduling mechanism is used, with priority 1.
- The default IP traffic profile, namely IP traffic profile 5 is used for rate limitation on a GPON port, with CIR of 2048 kbit/s.
- DBA profile 10 of Type4 is used and the maximum bandwidth in the ONT upstream direction is 100 Mbit/s.

To perform the preceding configurations, do as follows:

```
huawei(config)#dba-profile add profile-id 10 type4 max 102400
huawei(config)#ont-lineprofile gpon profile-id 5
huawei(config-gpon-lineprofile-5)#tcont 1 dba-profile-id 10
huawei(config-gpon-lineprofile-5)#qos-mode Priority-queue
huawei(config-gpon-lineprofile-5)#gem add 1 eth tcont 1 priority-queue 1
huawei(config-gpon-lineprofile-5)#mapping-mode vlan
huawei(config-gpon-lineprofile-5)#gem mapping 1 2 vlan 10
huawei(config-gpon-lineprofile-5)#commit
huawei(config-gpon-lineprofile-5)#quit
huawei(config-if-gpon-0/2)#ont confirm 1 ontid 1 sn-auth 32303131B39FD641
snmp ont-lineprofile-id 5
huawei(config-if-gpon-0/2)#quit
```

```
huawei(config)#service-port 101 vlan 100 gpon 0/2/1 ont 1 gemport 1 rx-cttr 5 tx-cttr
5
```

14.8 Congestion avoidance

14.8.1 Introduction

Definition

Congestion avoidance refers to a traffic control mechanism that monitors the utilization of network resources (such as queues or buffers), and drops packets to adjust the network traffic when congestion occurs, avoiding traffic overload on the network.

Congestion avoidance solves the issue of how to place packets in queue and how to drop packets.

Purpose

It uses some algorithms to avoid worse congestion and utilizes network bandwidth.

14.8.2 Basic Concepts

Congestion avoidance is implemented using packet drop algorithms. The MA5600T/MA5603T/MA5608T supports the following drop algorithms:

- Tail drop
- Priority-based early drop
- Color-based early drop

The following table explains these concepts.

Table 14-12 Basic concepts

Concept	Description	Effective Time
Tail drop	When a port queue is fully filled (to its maximum depth), the newly arriving packets are dropped until the queue has enough space to accept incoming traffic. Tail drop cannot ensure effective transmission of important data streams.	When the port queue is fully filled
Priority-based early drop	Packets with different priorities can be configured with different drop thresholds. Specifically, packets with higher priorities are configured with higher drop thresholds, while packets with lower priorities are configured with lower drop thresholds. When traffic congestion occurs on a port but the port queue is not fully filled, packets with higher priorities can enter queues that have a greater depth and are more burst-tolerant. In this way, these packets are less	When the port queue is partially filled

Concept	Description	Effective Time
	likely to be dropped. While packets with lower priorities are dropped preferentially because of the smaller queue depth.	
Color-based early drop	The system assigns different drop thresholds to different packets by marking the packets with different colors (yellow, green, or red; red packets are dropped directly) using the two rate three-color marker (trTCM) algorithm. In this way, when traffic congestion occurs on a port but the port queue is not fully filled, packets that do not exceed CIR (such packets are marked green) and packets that exceed CIR but do not exceed PIR (such packets are marked yellow) are forwarded; packets that exceed PIR (such packets are marked red) are dropped.	When the port queue is partially filled

14.8.3 Implementation Principle

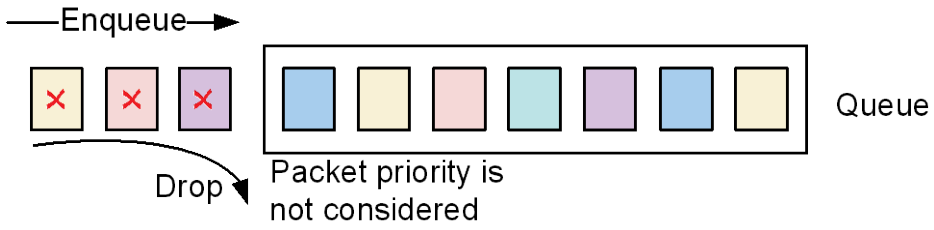
The MA5600T/MA5603T/MA5608T supports the following congestion avoidance algorithms:

- Tail drop
- Priority-based early drop
- Color-based early drop

Tail Drop

When a port queue is fully filled (to its maximum depth), the newly arriving packets are dropped until the queue has enough space to accept incoming traffic, as shown in Figure 14-30.

Figure 14-30 Tail drop



Tail drop applies to all queues and cannot be manually configured.

Downstream traffic is not dropped equally for different users in the same priority, because all the users are sharing the same queue in one PON port (outband direction) for each priority. In the same queue, the system does not recognize the packets of the different users, as the algorithm is drop-tail. When the output queue is full and tail drop is in effect, packets are dropped until the congestion is eliminated and the queue is no longer full. When the peer device sends such regular streams, the queue will fill during periods of congestion, it will

make the drop unequally for different users. When the traffic model is not regular streams, for example the traffic model in the Internet, the streams sending randomly, discards will become uniform, that the problem has no effect on the existing network.

Priority-based Early Drop

Packets with different priorities can be configured with different drop thresholds. Specifically, packets with higher priorities are configured with higher drop thresholds, while packets with lower priorities are configured with lower drop thresholds. When traffic congestion occurs on a port but the port queue is not fully filled, packets with higher priorities can enter queues that have a greater depth and are more burst-tolerant. In this way, these packets are less likely to be dropped. While packets with lower priorities are dropped preferentially because of the smaller queue depth.

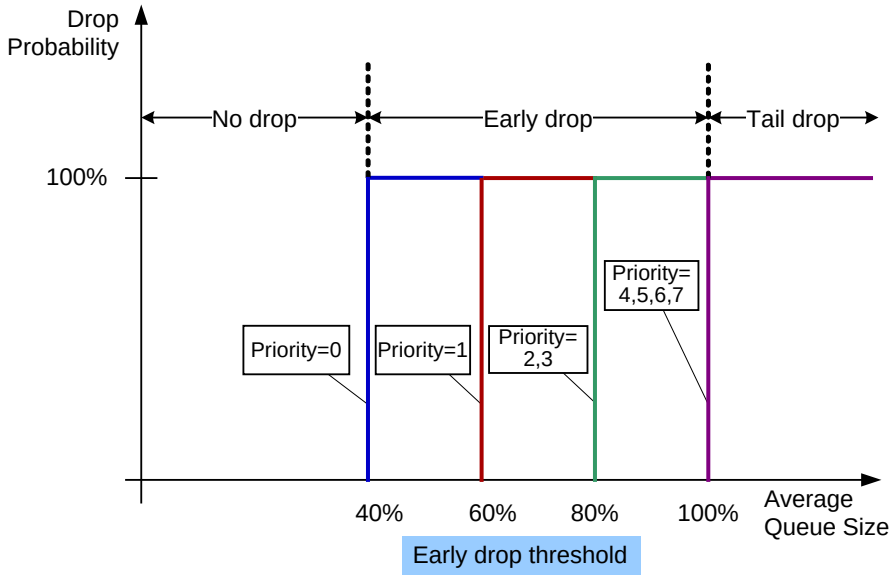
Priority-based early drop applies to scenarios in which packets with different priorities enter the same queue.

The **early-drop** command can be used to configure the early drop thresholds for packets of different priorities.

As shown in Figure 14-31, assume that packets with different priorities enter the same queue, the early drop thresholds are set to 40% and 60% for 0-priority and 1-priority packets, respectively, to 80% for 2-priority and 3-priority packets, and to 100% for 4-, 5-, 6-, and 7-priority packets.

- When the packets in the queue are less than 40% of the queue depth, all subsequent packets with different priorities can be enqueued.
- When packets in the queue increase to or exceed 40% of the queue depth, subsequent 0-priority packets are dropped while packets with higher priorities are enqueued.
- When packets in the queue increase to or exceed 60% of the queue depth, subsequent 1-priority packets are also dropped while packets with higher priorities are enqueued.
- When packets in the queue increase to or exceed 80% of the queue depth, subsequent 2-priority and 3-priority packets are also dropped, while packets with higher priorities (4, 5, 6, and 7) are enqueued.
- When packets in the queue increase to 100% of the queue depth, tail drop occurs. In this case, all subsequent packets are dropped.

Figure 14-31 Priority-based early drop

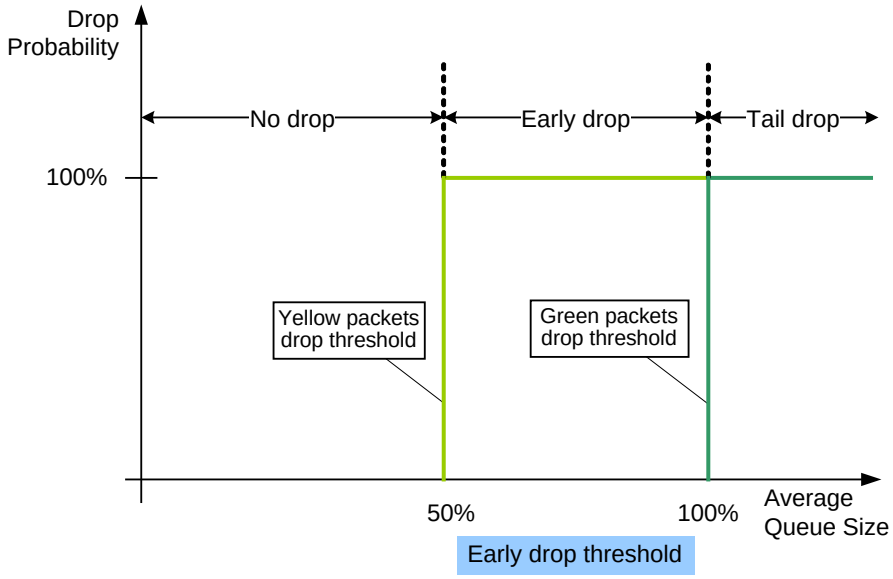


Color-based Early Drop

The system assigns different drop thresholds to different packets by marking the packets with different colors (yellow, green, or red; red packets are dropped directly) using the two rate three-color marker (trTCM) algorithm. In this way, when traffic congestion occurs on a port but the port queue is not fully filled, packets that do not exceed CIR (such packets are marked green) and packets that exceed CIR but do not exceed PIR (such packets are marked yellow) are forwarded; packets that exceed PIR (such packets are marked red) are dropped.

As shown in Figure 14-32, the drop threshold is 50% for yellow packets and 100% for green packets. The drop thresholds cannot be modified manually. If more than 50% of a queue is occupied, subsequent yellow packets cannot enter the queue but green packets can. When the queue is 100% occupied, green packets are also dropped.

Figure 14-32 Color-based early drop



In addition, for the OPGD/OPGE board, users can run the **early-drop color yellow** command to set the drop threshold for yellow packets with different priorities. Early drop for yellow packets is implemented in a similar way to priority-based early drop shown in Figure 14-31. The difference is that the former applies only to yellow packets with different priorities.

Color-based Early Drop - Weighted Random Early Discard (WRED)

The system assigns different drop thresholds to different packets by marking the packets with different colors (yellow, green, or red; red packets are dropped directly) using the trTCM algorithm. In this way, when traffic congestion occurs on a port but the port queue is not fully filled, packets that do not exceed CIR (such packets are marked green) and packets that exceed CIR but do not exceed PIR (such packets are marked yellow) are forwarded; packets that exceed PIR (such packets are marked red) are dropped.

The **wred-profile** command can be used to set the low drop limit (**low-limit**) and high drop limit (**high-limit**) for yellow and green packets, and the drop probability (**discard-probability**) for red packets.

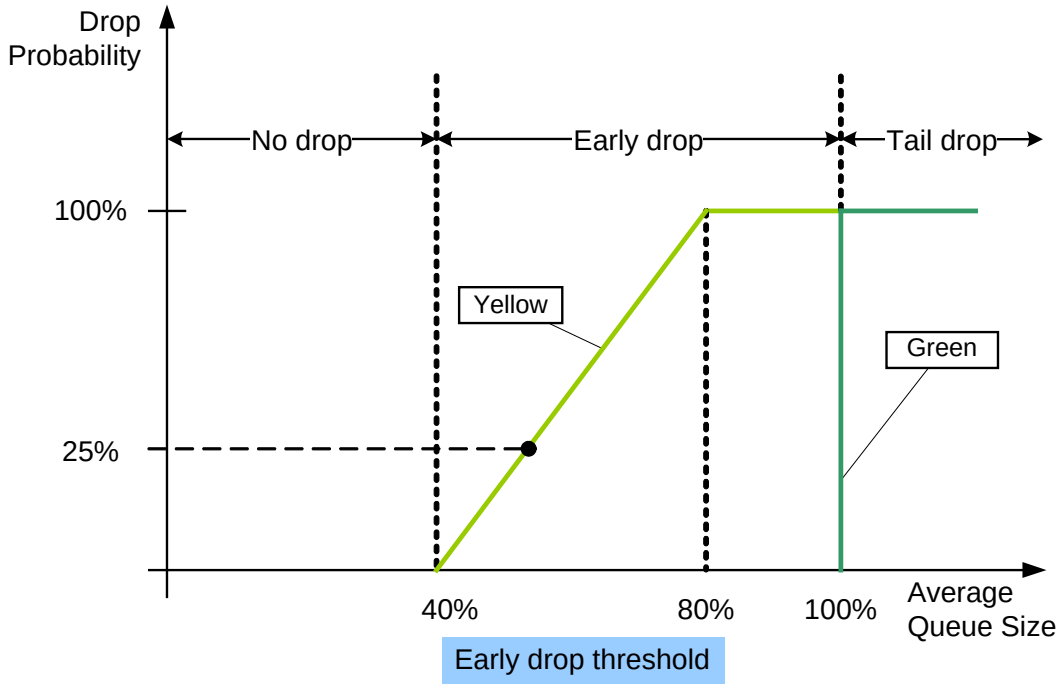
The drop probability at a specific time point derives from the following formula: Packet drop probability = (Usage of the packet buffer area - Low drop limit) ÷ (High drop limit - Low drop limit) x Drop probability of the packets reaching the high drop limit.

As shown in Figure 14-33, run the **wred-profile** command to specify the low drop limit to 40%, high drop limit to 80%, and drop probability to 100% for yellow packets, and no drop for green packets.

- When yellow packets in the queue are less than 40% of the queue depth, all subsequent yellow and green packets are enqueued.
- When yellow packets in the queue increase to or exceed 40% of the queue depth, all subsequent green packets are enqueued while subsequent yellow packets are dropped based on the drop probability specified for this time point (for example, 25%).
- When yellow packets in the queue increase to or exceed 80% of the queue depth, all subsequent yellow packets are directly dropped, while all subsequent green packets are enqueued.

- When packets in the queue increase to 100% of the queue depth (the queue is fully filled), tail drop occurs. In this case, all subsequent packets are dropped.

Figure 14-33 Color-based early drop - WRED



NOTE
In the figure above, the yellow line represents the drop probability curve.

14.8.4 Configuring the Congestion Avoidance

Context

- Congestion avoidance is implemented using packet drop algorithms, which include tail drop, priority-based early drop, and color-based early drop.
- The MA5600T/MA5603T/MA5608T can differentiate the services with different priorities in the same queue. The packet priority serves as a criterion for dropping packets.
- The MA5600T/MA5603T/MA5608T can also implement early drop based on the color of packets. When congestion occurs, the yellow packets are dropped.
- priority-based early drop and color-based early drop are mutually exclusive. Only one mode can be selected.

Procedure

- Configure priority-based early drop.
  a. Configure the early drop mode.

     Run the **early-drop mode pri-base** command to configure the priority-based early drop. After the configuration is completed, the system performs early drop according to the outer 802.1p priorities of the packets. When congestion occurs in a

queue, the packets are dropped according to the early drop thresholds of the priorities.

b. Configure the early drop threshold.

Run the **early-drop** command to configure the mapping between service priorities and drop thresholds. After configuration is successful, if the packets of the specified service priority reach the threshold of the queue (the percentage of the queue depth), subsequent packets of the same service priority will be dropped instead of entering the queue.

c. Query the configured early drop threshold.

Run the **display early-drop** command to query the configured early drop threshold.

- Configure color-based early drop.

a. Configure the early drop mode.

Run the **early-drop mode color-base** command to configure the color-based early drop.

According to the *CIR* and *PIR* parameters in the IP traffic profile, the system marks packets with colors. The packets within the CIR bandwidth are marked as green, and the packets between the CIR and PIR bandwidth are marked as yellow.

After the configuration is completed, green packets are allowed to pass, yellow packets that do not exceed the bandwidth can also pass, and yellow packets that exceed the bandwidth are dropped.

b. (For SPUF only) Configure the WRED early-discard based on packet color.

Run the **early-drop mode color-wred** command to configure the WRED early-discard based on packet color. After the configuration is successful, a WRED profile can bind to a packet queue. Then the packet color-based early-discard can be implemented according to the parameter settings in the WRED profile.

c. (For OPGD only) Configure the early drop threshold of yellow packets.

Run the **early-drop color yellow** command to configure the mapping between service priorities of yellow packets and drop thresholds. After the configuration is successful, if the yellow packets of the specified service priority reach the threshold of the queue (the percentage of the queue depth), subsequent yellow packets will be discarded instead of entering the queue.

d. Query the configured early drop threshold.

Run the **display early-drop** command to query the configured early drop threshold.

----End

Example

To configure the priority-based early drop, where,

- The early drop thresholds are set to 40% and 60% for 0-priority and 1-priority packets.
- The early drop thresholds are set to 80% for 2-priority and 3-priority packets.
- The early drop thresholds are set to 100% for 4-, 5-, 6-, and 7-priority packets.

Do as follows.

```
huawei(config)#early-drop mode pri-base
huawei(config)#early-drop cos0 40 cos1 60 cos2 80 cos3 80
{ <cr>|cos4<K>|cos5<K>|cos6<K>|cos7<K> }:
```

```
  Command:
        early-drop cos0 40 cos1 60 cos2 80 cos3 80
huawei(config)#display early-drop
{ <cr>|color<K>|mode<K> }:

  Command:
        display early-drop
  Early-drop is pri-base
  Pri-base early-drop config:
  --------------------------------
  Priority            Threshold
  --------------------------------
        0                  40
        1                  60
        2                  80
        3                  80
        4                 100
        5                 100
        6                 100
        7                 100
  --------------------------------
  Color-base early-drop config:
  --------------------------------
  Queue ID    WRED Profile Index
  --------------------------------
        0                   -
        1                   -
        2                   -
        3                   -
        4                   -
        5                   -
        6                   -
        7                   -
  --------------------------------
```

To configure the color-based early drop for OPGD board, where,

- The early drop thresholds are set to 40% and 60% for 0-priority and 1-priority packets.
- The early drop thresholds are set to 80% for 2-priority and 3-priority packets.
- The early drop thresholds are set to 100% for 4-, 5-, 6-, and 7-priority packets.

Do as follows.

```
huawei(config)#early-drop mode color-base
huawei(config)#early-drop color yellow cos0 40 cos1 60 cos2 80 cos3 80
{ <cr>|cos4<K>|cos5<K>|cos6<K>|cos7<K> }:

  Command:
        early-drop color yellow cos0 40 cos1 60 cos2 80 cos3 80
huawei(config)#display early-drop
{ <cr>|color<K>|mode<K> }:

  Command:
        display early-drop
  Early-drop is color-base
  Pri-base early-drop config:
```

```
--------------------------------
Priority           Threshold
--------------------------------
      0                 40
      1                 60
      2                 80
      3                 80
      4                100
      5                100
      6                100
      7                100
--------------------------------
Color-base early-drop config:
--------------------------------
Queue ID    WRED Profile Index
--------------------------------
      0                  -
      1                  -
      2                  -
      3                  -
      4                  -
      5                  -
      6                  -
      7                  -
--------------------------------
```

14.9 Congestion Management

14.9.1 Introduction

Definition

If packets arrive faster than they are forwarded on a port, traffic congestion occurs on the port. Congestion management is used to manage and control traffic congestion.

Congestion management is implemented using queuing techniques for transmitting packets out of a queue.

Purpose

When traffic congestion occurs on an outbound interface, a proper queue scheduling mechanism guarantees required QoS parameters (such as bandwidth, latency, and jitter) for a certain type of packets.

14.9.2 Basic Concepts

Congestion management is implemented using queuing techniques. The MA5600T/MA5603T/MA5608T supports the following queuing techniques:

- Priority queuing (PQ)
- Weighted round robin (WRR)
- PQ+WRR

The following table explains the basic concepts of congestion management.

Table 14-13 Basic concepts

Concept	Description	Remarks
PQ	In PQ, the strict priority scheduling algorithm is used. In this scheduling algorithm, eight priorities are defined. Packets with higher priorities are scheduled first.	Packets of important services are processed preferentially.
WRR	In the WRR scheduling algorithm, packets are scheduled based on the assigned weights.	Specific QoS guarantees are provided for each queue.
PQ+WRR	The combination of PQ+WRR ensures that some of the packets with high priorities are scheduled first and the remaining packets are scheduled based on the specified weights.	This mechanism ensures effective transmission of services with high priorities. In the mean time, services with low priorities are scheduled in a timely manner when there is available bandwidth.
Enqueuing priority	The enqueuing priority of a packet determines the queue that the packet will enter.	There are several sources of the enqueuing priority for a packet and the priority value ranges from 0 to 7.
Queue buffering	The queue depth determines the queue's capability of processing burst packets. A larger buffering space means a better capability of processing burst packets and a lower chance to lose packets, but a larger delay in processing packets.	-

14.9.3 Implementation Principle

The MA5600T/MA5603T/MA5608T processes packets as follows before packets enter queues or when they are in queues:

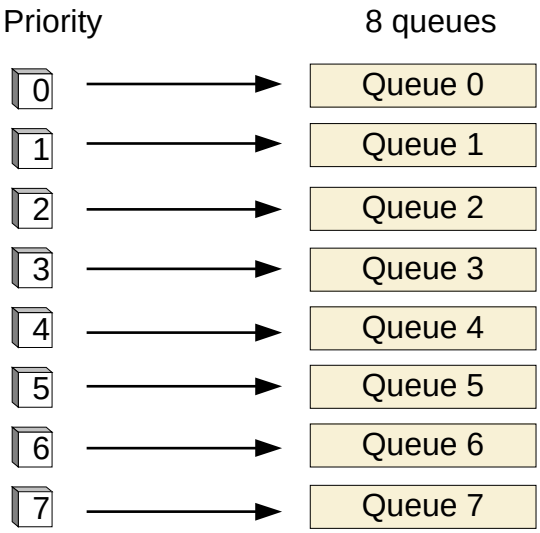
1. Traffic classification: For details, see the 14.5.2 Implementation Principle.
2. Enqueuing: Packets enter different queues based on the mapping between packet priorities and queues.
3. Scheduling: The MA5600T/MA5603T/MA5608T schedules packets using a specific scheduling algorithm after the packets are enqueued.

Enqueuing

After priority processing, packets enter different queues based on the mapping between packet priorities and queues. By default, the mapping between packet priorities and queues is constant, as shown in Figure 14-34.

The larger the queue ID, the higher the forwarding priority for packets in the queue. Among all the eight queues, queue 7 has the highest priority.

Figure 14-34 Default mapping between packet priorities and queues



The system also supports a flexible mapping between packet priorities and queues. You can run the **cos-queue-map** command to map the 802.1p priority of the packet to any queue. That is, a queue can contain multiple 802.1p priorities or no 802.1p priorities.

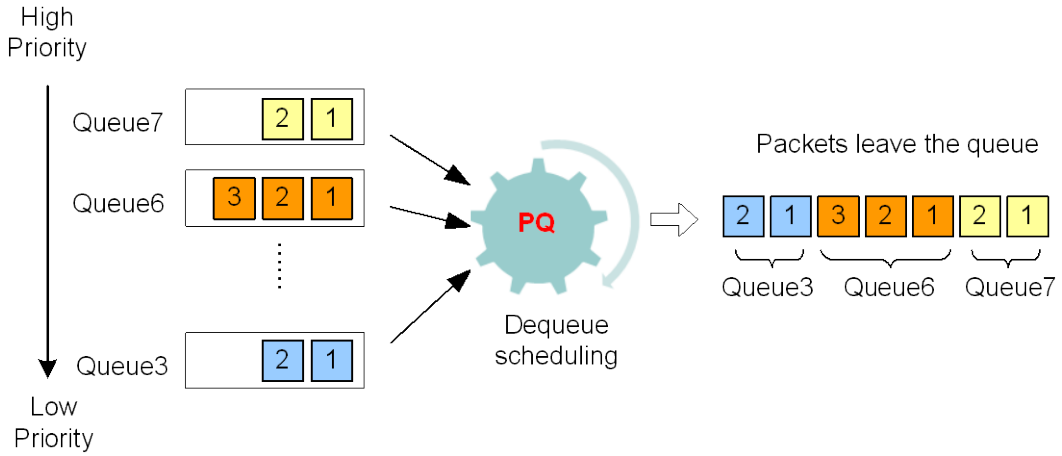
The configuration of the mapping between packet priorities and queues takes effect globally. Generally, use the default value in the system and do not modify the parameters unless you have specific requirements.

Queue Scheduling: PQ

Priority queuing (PQ) queues are classified into high-priority queues, medium-priority queues, normal-priority queues, and low-priority queues in the descending priority order. As shown in Figure 14-35, PQ allows the packets in a high-priority queue to exit the queue and be transmitted by the MA5600T/MA5603T/MA5608T. After such a transmission is completed, PQ performs the same on all packets in a medium-priority queue, a normal-priority queue, and then a low-priority queue one by one.

In this way, packets in a queue with a higher priority precede packets in a queue with a lower priority and therefore are processed preferentially, even in case of congestion. This mechanism ensures that packets for critical services are processed first. Packets of non-critical services (such as email service) are processed only when the network has sufficient resources after critical services have been processed, thereby utilizing network resources efficiently.

Figure 14-35 PQ scheduling mechanism

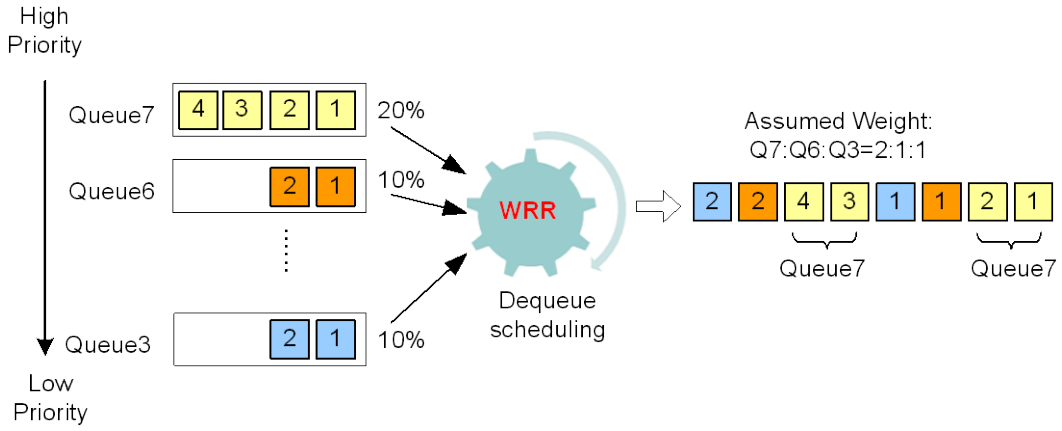


Queue Scheduling: WRR

Weighted round robin (WRR) classifies packets and places packets into the corresponding queues according to the packet classification results. As shown in Figure 14-36, WRR queues are assigned bandwidth on a port according to the bandwidth percentages defined by the user. When packets are waiting to exit the queues, WRR takes a certain number of packets from the queue and transmits them from the port according to the pre-defined bandwidth percentage.

In WRR scheduling mode, the queues are scheduled in turn based on certain weight values. This mechanism ensures that each queue can be scheduled. When a queue is empty, the next queue is scheduled immediately. In this way, the bandwidth resources can be fully utilized.

Figure 14-36 WRR scheduling mechanism

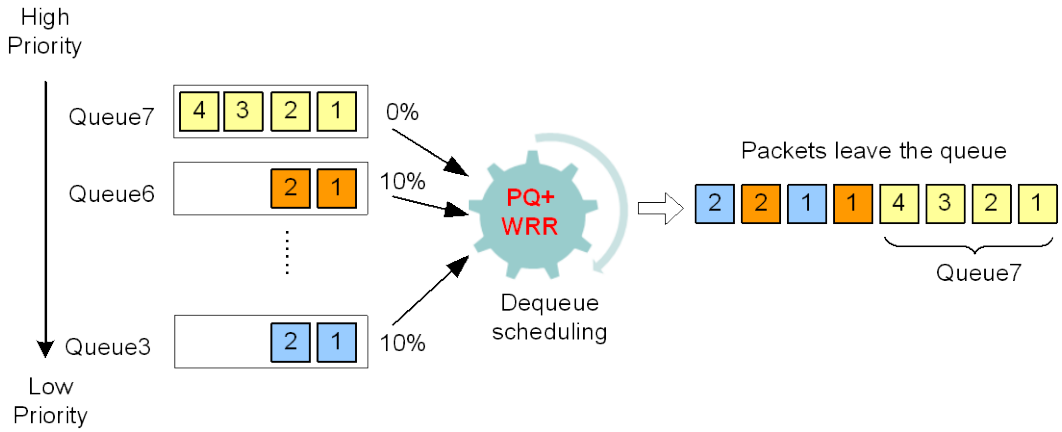


Queue Scheduling: PQ+WRR

PQ+WRR is a combination of the PQ scheduling mode and WRR scheduling mode. When the weight value of a queue is 0, the queue scheduling mode is PQ+WRR. As shown in Figure 14-37, in this mode, the system schedules the queues with the weight value 0 in the PQ mode, and then schedules other queues in the WRR mode.

This scheduling mode is flexible. The services that must be guaranteed are scheduled in PQ mode, and the services with lower priorities are scheduled in WRR mode when there is available bandwidth. In this way, services with higher priorities are ensured and those with lower priorities can obtain bandwidth whenever there is available bandwidth.

Figure 14-37 PQ+WRR scheduling mechanism



14.9.4 Configuring the Congestion Management

Context

Congestion management contains the following content:

- Mapping between the queue and the 802.1p priority
- Queue depth
- Queue scheduling mode

And, the ultimate effect of the congestion management is determined by all of them.

Procedure

Configure the mapping between the queue and the 802.1p priority.

The configuration of the mapping between priorities and queues takes effect globally. Generally, use the default value in the system and do not customize the parameters unless you have specified requirements.

After the configuration, packets with different 802.1p priorities are mapped to the specified queues based on the configured mapping. This enhances the flexibility of mapping packets to queues.

1. Run the **cos-queue-map** command to configure the mapping between the 802.1p priority and the queue.
2. Run the **display cos-queue-map** command to query the mapping between the 802.1p priority and the queue.

Step 1 (Optional) Configure the queue depth.

By default, the queue depth of queue 0-7 is 13%, 6%, 13%, 13%, 12%, 12%, 25% and 6% respectively. Do not modify the value unless for special requirement.

configure the queue depth (the queue buffer space) to re-allocate buffer space to the queues, therefore to improve the flexibility of QoS.

1. Run the **queue-buffer** command to set the buffer size percentage for packet queues of the service boards in the system.
2. Run the **display queue-buffer** command to query the buffer size percentage for queues of the service boards in the system.

Step 2 Configure the queue scheduling mode.

By default, the PQ mode is used. Properly set according to your actual requirements.

1. Run the **queue-scheduler** command to configure the queue scheduling mode.

NOTE

Range of *queue0-weight*: 0-100, 255, where 0 indicates that the strict PQ scheduling mode is used and 255 indicates that the queue is not used. The sum of the weight values of all queues must be 0 or 100.

2. Run the **display queue-scheduler** command to query the configuration of the queue scheduling mode.

----End

Example

Configure the queue scheduling mode to 3PQ+5WRR. Where,

- Queues 5-7 adopt the PQ mode
- Queues 0-4 adopt the WRR mode
- Queues 0-4 with the weights of 10, 10, 20, 20, and 40 respectively
- Other parameters use default settings

After that, packets with priority 5, 6,or 7 are scheduled in the PQ mode, and packets with other priority are scheduled in the WRR mode.

```
huawei(config)#queue-scheduler wrr
{ queue0-weight<U><0,255> }:10
{ queue1-weight<U><0,255> }:30
{ queue2-weight<U><0,255> }:20
{ queue3-weight<U><0,255> }:20
{ queue4-weight<U><0,255> }:20
{ queue5-weight<U><0,255> }:0
{ queue6-weight<U><0,255> }:0
{ <cr>|queue7-weight<U><0,255> }:0

 Command:
       queue-scheduler wrr 10 30 20 20 20 0 0 0

huawei(config)#display queue-scheduler
 Queue scheduler mode : WRR
 --------------------------------
 Queue  Scheduler Mode  WRR Weight
 --------------------------------
    0  WRR                 10
    1  WRR                 30
```

```
    2  WRR                 20
    3  WRR                 20
    4  WRR                 20
    5  PQ                  --
    6  PQ                  --
    7  PQ                  --
--------------------------------
```

14.10 ACL

Using the preset access control list (ACL) policy, the system permits or refuses data packets to pass.

14.10.1 Overview

Definition

The access control list (ACL) policy defines a series of matching rules, according to which the packets to be filtered are identified. The packets identified are permitted or refused to pass according to the preset rules.

ACL-based traffic filtering is a prerequisite for quality of service (QoS). ACL together with QoS improves system security.

Benefits

Mutual access between internal networks and communication between internal and external networks are primary requirements for enterprise networks. To ensure internal network security, a security policy is required to allow unauthorized users to access specified network resources so as to control access. With the ACL, network traffic can be filtered and network access is controlled.

After being bound to a QoS operation, the ACL helps to implement the following functions:

- Limiting network traffic and improving network performance
- Ensuring safe network access
- Determining which type of communication traffic to be forwarded or blocked on a port on the network device, such as a router or switch

Controlling packets on an access port when the ACL is used for ACL-based Firewall Filtering, prohibiting unauthorized users from logging in to the system and ensuring device safety

14.10.2 Basic Concepts

ACL Type

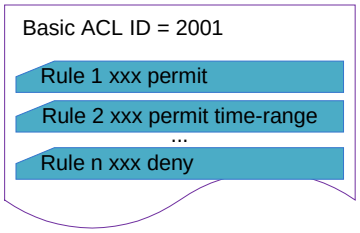
The system supports IPv4 and IPv6 ACLs, that is, ACLv4 and ACLv6. ACLs include basic ACLs, advanced ACLs, link-layer ACLs, and user-defined ACLs.

Table 14-14 ACL types

Type	Serial Number Range	Feature
Basic ACL	2000-2999	ACLv4: The rules of a basic ACL can be defined only according to the Layer 3 source IP address and the fragment field, for analyzing and processing data packets. ACLv6: The rules of a basic ACL can be defined only according to the Layer 3 source IP address and subnet prefix length, for analyzing and processing data packets.
Advanced ACL	3000-3999	Compared with a basic ACL, an advanced ACL allows for a wider scope of more accurate and flexible definition of the rules according to the following data packet information: • Source address • Destination address • IP bearer protocol types 0-255 (GRE, ICMP/ICMPv6, IP/IPv6, IPinIP, TCP, IPv6-ah, IPv6-esp, and OSPF) • TCP source port • TCP destination port • ICMP/ICMPv6 protocol type • ICMP/ICMPv6 code • Priority TOS/IP precedence/DSCP
Link-layer ACL	4000-4999	The rules of a link-layer ACL can be defined according to the following information: • Source MAC address • VLAN ID • Layer 2 protocol type • Destination MAC address • 802.1p priority
User-defined ACL	5000-5999	The rules of a user-defined ACL can be defined according to any 32 bytes of the first 80 bytes in a Layer 2 frame. • IPoE matching: matches packets whose Ethernet packet header is IPv4-encapsulated, including untagged IP packets, one-tagged IP packets, and two-tagged IP packets. • Non-IPoE matching: matches IPv4 packets whose Ethernet packet header is not IP-encapsulated, that is non-IPoE IPv4 packets, including untagged, one-tagged, and two-tagged non-IPoE packets, or multi-tagged packets.

Rule

Rules are the main body of an ACL. An ACL can be composed of multiple rules, as shown in the following figure.



A rule has the following characteristics:

- Each rule belongs to an absolute ACL and is of the same type as the ACL.
- Each rule corresponds to a permit or deny action.
- Each rule can be configured with a time range in which the ACL takes effect.

14.10.3 ACL Rule Matching Sequence

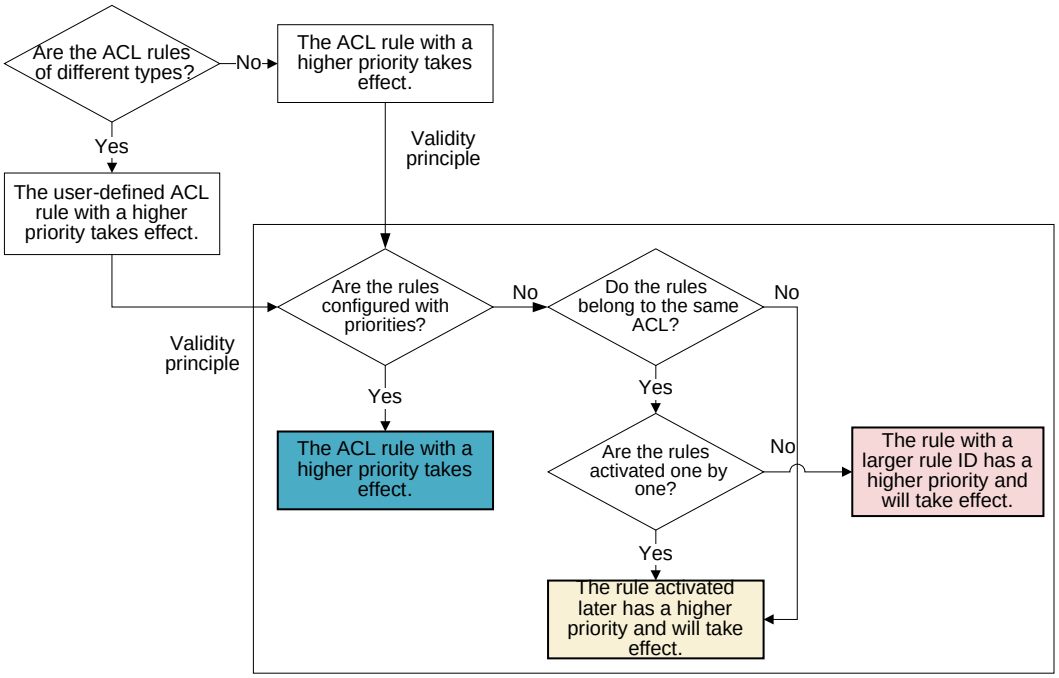
An ACL rule needs to be bound to a QoS policy so that it can take effect on a port. Two general ACL matching principles are provided as follows:

- An ACL rule is valid only when it is within the period of *time-range-name*.
- ACL rules are matched based on their priorities in a descending order; this process stops once a rule is matched.

Packets Matching Two or More ACL Rules

When packets reach a port, the system matches the packets with two or more ACL rules, as shown in Figure 14-38.

Figure 14-38 Validity of ACL rules when packets match multiple ACL rules



When packets match two or more ACL rules:

- If ACL rules are of different types, the priority of a user-defined rule is higher than that of all non-user-defined rules. If the user-defined rule is used, the other rules may be invalid. Therefore, exercise caution when using this rule.
- If ACL rules are of the same type but configured with different priorities, the ACL rule of a higher priority will be valid.
- If ACL rules are of the same type and no priority is configured for these ACL rules, do as follows:
  - For rules of the same ACL, if they are activated at the same time, the rule with a larger *rule-id* has a higher priority.
  - For rules of the same ACL, if they are activated one by one, the rule activated later has a higher priority than the one activated earlier.
  - For rules belonging to different ACLs, the rule activated later has a higher priority than the one activated earlier.

In other cases:

- When both the Layer 3 ACL (basic ACL and advanced ACL) and the Layer 2 ACL (link-layer ACL) are issued, all rules use the priority that is configured for the Layer 2 ACL rule.
- When both the IPv6-based rule and the link-layer rule exist, the link-layer rule prevails even if a higher priority is configured for the IPv6-based rule.
- Among the rules issued to the routing interface or firewall, the rule with a smaller *rule-id*has a higher priority, regardless of the activation sequence or the configured priority. The rules are used to match the packets based on **rule-id** in an ascending order. Once the rule with a smaller **rule-id** matches the packets, the rule matching stops, that is, other rules with a larger **rule-id** are invalid.

- When you run the **packet-filter** command to use an ACL and specify the **to-cpu** packet, the rules are matched based on priorities in a descending order; this process stops once a rule is matched. The matching sequence is irrelative to whether the rule is a user-defined rule, an IPv6 rule, or a rule configured with a priority. Rules are matched based on the following principles only:
  - For rules of the same ACL, if they are activated at the same time, the rule with a larger *rule-id* has a higher priority.
  - For rules of the same ACL, if they are activated one by one, the rule activated later has a higher priority than the one activated earlier.
  - For rules belonging to different ACLs, if they are issued to the port from different ACLs, the rule activated later has a higher priority than the one activated earlier.

Packets Matching No ACL Rule

When no packet matches any rule in the ACL, the processing is as follows:

- For the ACL referenced by route interfaces or firewalls, if no packet matches any rule in the ACL, the traffic behavior will be performed, that is, the default operation is permit. You can run the **firewall default** command to configure the default operation for the firewall to **permit** or **deny**.

  For example, when you run the **firewall packet-filter** command to filter out the packets that are about to pass the firewall, and the packet whose source IP address is 10.10.10.10/24 does not match any rule in the ACL, the packet is allowed to pass the firewall.
- For the ACL referenced by common ports, if no packet matches any rule in the ACL, no traffic behavior will be performed.

  For example, when you run the **packet-filter** command to filter out the packets, and the packet whose source IP address is 10.10.10.10/24 does not match any rule in the ACL, the device will receive this packet. When you run the **traffic-limit** command to limit the traffic of packets that match the ACL rule and the packet whose source IP address is 10.10.10.10/24 does not match any rule in the ACL, the device only forwards the packet and does not perform the traffic limiting.

14.10.4 ACL Rule Matching Process

ACL Rule Matching Process

The system will match the inbound packets according to the defined ACL rules:

- If the packets match an ACL rule, they are performed with further QoS actions, including packet filtering, priority marking, rate limiting, traffic statistics measurement, packet redirection, and packet mirroring. After being processed using the preceding QoS actions, the packets are forwarded in the outbound direction.
- If the packets do not match an ACL rule, the packets are directly forwarded.

NOTE

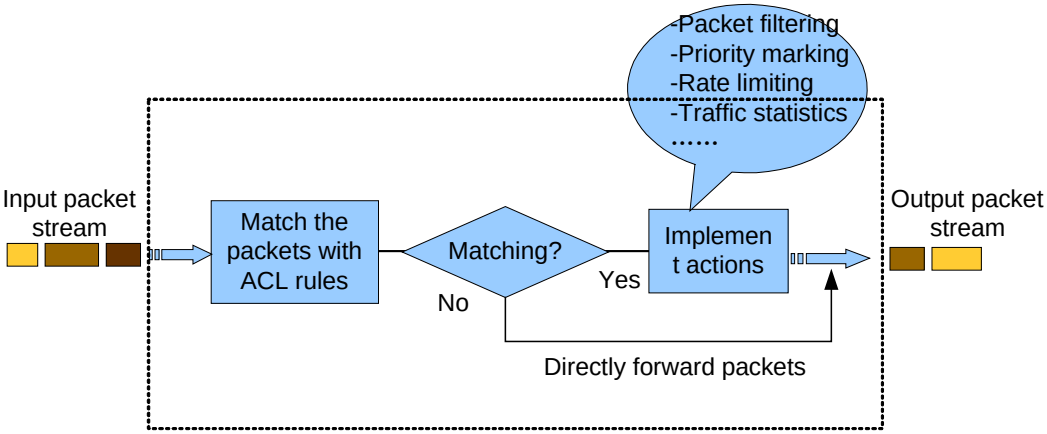
However, when the ACL rule is applied to the packet filtering firewall, the packets that do no match the ACL rule are processed according to the default firewall policy (run the **firewall default { permit | deny }** command to configure the default firewall policy). If the firewall default policy is deny, the packets will be discarded.

The following table lists the ACL-based QoS policies and relevant functions and configuration methods.

ACL-based QoS Policy	Function	Configuration Method
Packet filtering	Determines whether to discard packets or allow the packets to pass according to the matching result of the ACL rule (deny or permit). Only this policy supports the deny operation by an ACL rule.	Configuration command: **packet-filter**
Priority marking	Marks priorities for packets that match an ACL rule on a specified port or port list by type, such as ToS, DSCP, and 802.1p.	Configuration command: **traffic-priority**
Rate limiting	Limits the rate of packets that match an ACL rule on a specified port or port list. If traffic on a port exceeds the limit, the excessive packets will be discarded based on the trTCM algorithm or be configured with a new DSCP priority.	Configuration command: **traffic-limit**
Traffic statistics	Measures the packets that match an ACL rule on a specified port or port list in terms of packet number and byte.	Configuration command: **traffic-statistic**
Packet redirection	Redirects the packets that match an ACL rule on a specified port or port list to a new forwarding destination port (the original port is not used to forward packets that match an ACL rule).	Configuration command: **traffic-redirect**
Packet mirroring	Mirrors the packets that match an ACL rule on a specified port and copies the packets to other ports. If packets received or transmitted through a port need to be monitored during device maintenance or fault diagnosis, run the command to mirror packets to be monitored to other ports. **NOTE** Based on industry experience, the mirroring feature may involve obtaining personal data of users and the content of users' communications (the product does not save, parse, or process such information) for the purpose of safeguarding network operation and protecting services. Huawei alone is unable to collect or save the personal data of users and the content of users' communications. It is suggested that you activate the interception-related functions based on the applicable laws and regulations in terms of purpose and scope of usage. You are obligated to take considerable measures to ensure that the personal data of users and the content of users' communications are fully protected when the personal data and the content are being used and saved.	Configuration command: **traffic-mirror**

Figure 14-39 illustrates ACL-based packet filtering and processing.

Figure 14-39 ACL-based packet filtering and processing



14.10.5 Matching Principle for the User-defined ACL Rule

The user-defined ACL rule is matched using any 32 bytes of the first 80 bytes of a Layer 2 data frame. The following describes the format of the Layer 2 data frame.

Format of the Layer 2 Data Frame

The following uses a Layer 2 IPv4 data frame as an example.

Figure 14-40 First 64 bytes of an IPv4 data frame

Byte	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
	AA	AA	AA	AA	AA	AA	BB	BB	BB	BB	BB	BB	CC	CC	CC	CC
	DD	DD	EE	FF	GG	GG	HH	HH	II	II	JJ	KK	LL	LL	MM	MM
	MM	MM	NN	NN	NN	NN	OO	OO	PP	PP	QQ	QQ	QQ	QQ	RR	RR
	RR	RR	SS	TT	UU	UU	VV	VV	VV	VV	VV	VV	VV	VV	VV	VV

Byte	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
	AA	AA	AA	AA	AA	AA	BB	BB	BB	BB	BB	BB	CC	CC	CC	CC
	DD	DD	EE	FF	GG	GG	HH	HH	II	II	JJ	KK	LL	LL	MM	MM
	MM	MM	NN	NN	NN	NN	OO	OO	PP	PP	QQ	QQ	QQ	QQ	RR	RR
	RR	RR	SS	TT	UU	UU	VV	VV	VV	VV	VV	VV	VV	VV	VV	VV

shows the sequence of the first 64 bytes of the Layer 2 IPv4 data frame, in which:

- Different letters indicate different field values. For example, A refers to the destination MAC address and C refers to the Ethernet type+VLAN tag.
- The first letter of each letter group indicates the offset value of the field. For example, A indicates that the offset value is 0 and C indicates that the offset value is 12.

Table 14-15 lists the meaning and offset value of each letter.

Table 14-15 Meanings of letters and their offset values

Letter & Meaning	Offset Value	Letter & Meaning	Offset Value	Letter & Meaning	Offset Value
A: destination MAC address	0	I: flags	24	Q: serial number	42
B: source MAC address	6	J: time to live	26	R: acknowledgment field	46
C: Ethernet type+VLAN tag	12	K: protocol ID ("6" represents TCP and "17" represents UDP)	27	S: IP header length and reserved bit	50
D: protocol type	16	L: IP checksum	28	T: reserved bit and flags bit	51
E: IP version number	18	M: source IP address	30	U: window size	52
F: service type	19	N: destination IP address	34	V: Others	54
G: length of the IP packet	20	O: TCP source port	38	-	-
H: ID	22	P: TCP destination port	40	-	-

NOTE
The offset value of each field is the offset value in data frame ETH II+VLAN tag.

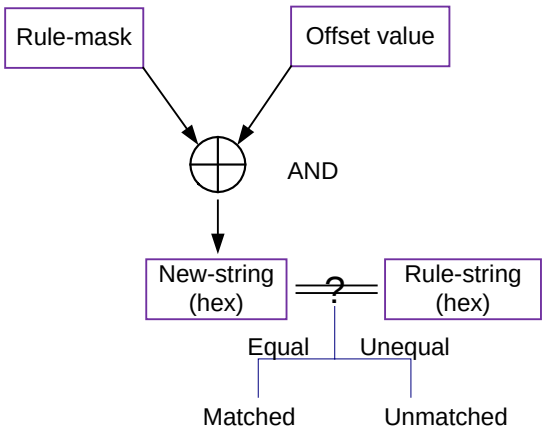
Matching Principle

The user-defined ACL rule involves the following key concepts:

- rule-string: indicates the character string of the user-defined ACL rule. The character string is in hexadecimal notation. The number of characters in the string must be an even number ranging from 2 to 8.
- rule-mask: indicates the mask of the user-defined ACL rule. It is used to perform the logical AND operation with specified fields in the data packets.
- offset: indicates the offset. With the header of the packet as the reference point, it specifies a certain field in the Layer 2 data frame.

With the user-defined ACL rule, users can use rule-mask and offset to extract any 32 bytes of the first 80 bytes of a Layer 2 data frame and compare them with the rule-string so as to find the matched data frame. Figure 14-41 shows the matching principle.

Figure 14-41 Matching principle for the user-defined ACL rule



14.10.6 Configuring Traffic Management Based on ACL Rules

The ACL can be used to implement flexible traffic classification according to user requirements. After traffic classification based on ACL rules is completed, you can perform QoS for the traffic streams.

Configuration Differences Between IPv4 ACLs and IPv6 ACLs

This topic describes differences regarding to configuration between IPv4 ACLs and IPv6 ACLs. It is recommended that you know well about how to configure IPv4 ACLs and then configure IPv6 ACLs based on their differences.

Configuration Differences

- The configuration differences between IPv4 ACLs and IPv6 ACLs are as follows:
  - IPv6 and IPv4 have different IP address formats and packet formats, so the **ipv6** parameter must be specified for configuring IPv6 basic ACLs and advanced ACLs. Use the **ipv6** parameter to choose between IPv4 ACLs and IPv6 ACLs.
  - IPv4 and IPv6 have the same link-layer packet encapsulation format, so configurations do not differentiate IPv6 link-layer ACLs and IPv4 link-layer ACLs.
  - Users define packets matching ACLs based on the packet type. IPv4 and IPv6 have the same packet command for user-defined ACLs, so configurations do not differentiate IPv6 user-defined ACLs and IPv4 user-defined ACLs. When user-defined ACLs are used for filtering packets, the protocol type of the packets must be the same as the protocol type of the ACL rules. If they are different, filtering may encounter errors.

Configuring an ACL Rule

Context

ACLs include:

- Basic ACL
- Advanced ACL
- Link-layer ACL
- User-defined ACL

For features of each type of ACL, see "14.10.2 Basic Concepts."

Precaution

Because the ACL is flexible in use, Huawei provides the following suggestions on its configuration:

- It is recommended that you define a general rule, such as permit any or deny any, in each ACL, so that each packet has a matching traffic rule that determines to forward or filter the unspecified packet.
- The activated ACL rules share the hardware resources with the protocol modules (such as Dynamic Host Configuration Protocol (DHCP) module and Internet Protocol over ATM (IPoA) module). In this case, the hardware resources are limited and may be insufficient. To prevent the failure to enable other service functions due to insufficient hardware resources, it is recommended that you enable the protocol module first and then activate ACL rules in the data configuration. If you fail to enable a protocol module, perform the following steps:
  a. Check whether ACL rules occupy too many resources.
  b. If ACL rules occupy too many resources, deactivate or delete the unimportant or temporarily unused ACL configurations, and then configure and enable the protocol module.

Procedure

- Configure a basic ACL rule.

  It is applicable to the scenario where the device needs to match packets by source IP address.

  a. (Optional) Set a time range.

     Run the **time-range** command to create a time range, which can be used when an ACL rule is created.

  b. Create a basic ACL.

     Run the **acl** command to create a basic ACL, and then enter the ACL mode. The serial number of a basic ACL ranges from 2000 to 2999.

  c. Configure a basic ACL rule.

     In the acl-basic mode, run the **rule** command to create a basic ACL rule. The parameters are as follows:

     - *rule-id*: indicates the ACL rule ID. To create an ACL rule with a specified ID, use this parameter.
     - **permit**: indicates the keyword for allowing the data packets that meet related conditions to pass.

  - **deny**: indicates the keyword for discarding the data packets that meet related conditions.
  - **time-range**: indicates the keyword of the time range during which the ACL rule will take effect.

- Configure an advanced ACL rule.

  It is applicable to the scenario where the device needs to match data packets by source IP address, destination IP address, type of protocol running over IP, and protocol feature, such as TCP source port, TCP destination port, and ICMP type of the data packets.

  a. (Optional) Set a time range.

     Run the **time-range** command to create a time range, which can be used when an ACL rule is created.

  b. Create an advanced ACL.

     Run the **acl** command to create an advanced ACL, and then enter the acl-adv mode. The serial number of an advancedACL ranges from 3000 to 3999.

  c. Configure an advanced ACL rule.

     In the acl-adv mode, run the **rule** command to create an ACL rule. The parameters are as follows:

     - *rule-id*: indicates the ACL rule ID. To create an ACL rule with a specified ID, use this parameter.
     - **permit**: indicates the keyword for allowing the data packets that meet related conditions to pass.
     - **deny**: indicates the keyword for discarding the data packets that meet related conditions.
     - **time-range**: indicates the keyword of the time range during which the ACL rule will take effect.

- Configure a link-layer ACL rule.

  It is applicable to the scenario where the device needs to match packets by link-layer information such as source MAC address, source VLAN ID, Layer 2 protocol type, and destination MAC address.

  a. (Optional) Set a time range.

     Run the **time-range** command to create a time range, which can be used when an ACL rule is created.

  b. Create a link-layer ACL.

     Run the **acl** command to create a link-layer ACL, and then enter the acl-link mode. The serial number of a link-layer ACL ranges from 4000 to 4999.

  c. Configure a link-layer ACL rule.

     In the acl-link mode, run the **rule** command to create a link-layer ACL rule. The parameters are as follows:

     - *rule-id*: indicates the ACL rule ID. To create an ACL rule with a specified ID, use this parameter.
     - **permit**: indicates the keyword for allowing the data packets that meet related conditions to pass.
     - **deny**: indicates the keyword for discarding the data packets that meet related conditions.
     - **time-range**: indicates the keyword of the time range during which the ACL rule will take effect.

- Configure a user-defined ACL rule.

It is applicable to the scenario where the device needs to match packets by any 32 bytes of the first 80 bytes of a Layer 2 data frame.

Configuring a user-defined ACL requires a deep understanding of the Layer 2 data frame structure. Be sure to make a data plan according to the format of the Layer 2 data frame. Refer to "14.10.5 Matching Principle for the User-defined ACL Rule" for details.

a. (Optional) Set a time range.

   Run the **time-range** command to create a time range, which can be used when an ACL rule is created.

b. Create a user-defined ACL.

   Run the **acl** command to create a user-defined ACL, and then enter the acl-user mode. The serial number of a user-defined ACL ranges from 5000 to 5999.

c. Configure a user-defined ACL rule.

   In the acl-user mode, run the **rule** command to create an ACL rule. The parameters are as follows:

   - *rule-id*: indicates the ACL rule ID. To create an ACL rule with a specified ID, use this parameter.
   - **permit**: indicates the keyword for allowing the data packets that meet related conditions to pass.
   - **deny**: indicates the keyword for discarding the data packets that meet related conditions.
   - *rule-string*: indicates the character string of the user-defined ACL rule. The character string is in hexadecimal notation. The number of characters in the string must be an even number.
   - *rule-mask*: indicates the mask of the user-defined ACL rule. It is a positive mask, used to perform the AND operation with the data packets for extracting the information from the data packets.
   - *offset*: indicates the offset. With the header of the packet as the reference point, it specifies the byte from which the AND operation begins. Together with the rule mask, it extracts a character string from the packets.
   - **ipoe**: indicates that the Ethernet packet header encapsulates an IP packet, including untagged, one-tagged, and two-tagged IP packets.
   - **non-ipoe**: indicates that the Ethernet packet header encapsulates a non-IP packet, including the untagged, one-tagged, and two-tagged non-IP packets, or multi-tagged packets.
   - **time-range**: indicates the keyword of the time range during which the ACL rule will take effect.

----End

Example

To configure port 0/2/0 on the MA5600T/MA5603T/MA5608T to receive only the packets from address 2.2.2.2 from 00:00 to 12:00 on Fridays, and to discard the packets from other addresses, do as follows:

```
huawei(config)#time-range time1 00:00 to 12:00 fri
huawei(config)#acl 2000
huawei(config-acl-basic-2000)#rule permit source 2.2.2.2 0.0.0.0 time-range time1
huawei(config-acl-basic-2000)#rule deny time-range time1
huawei(config-acl-basic-2000)#quit
```

```
huawei(config)#packet-filter inbound ip-group 2000 port 0/2/0
huawei(config)#save
```

Assuming that the service board of the MA5600T/MA5603T/MA5608T resides in slot 1 and belongs to a VLAN, and the IP address of the VLAN Layer 3 interface is 10.10.10.101, to prohibit the ICMP (such as ping) and telnet operations from the user side to the VLAN interface on the device, do as follows:

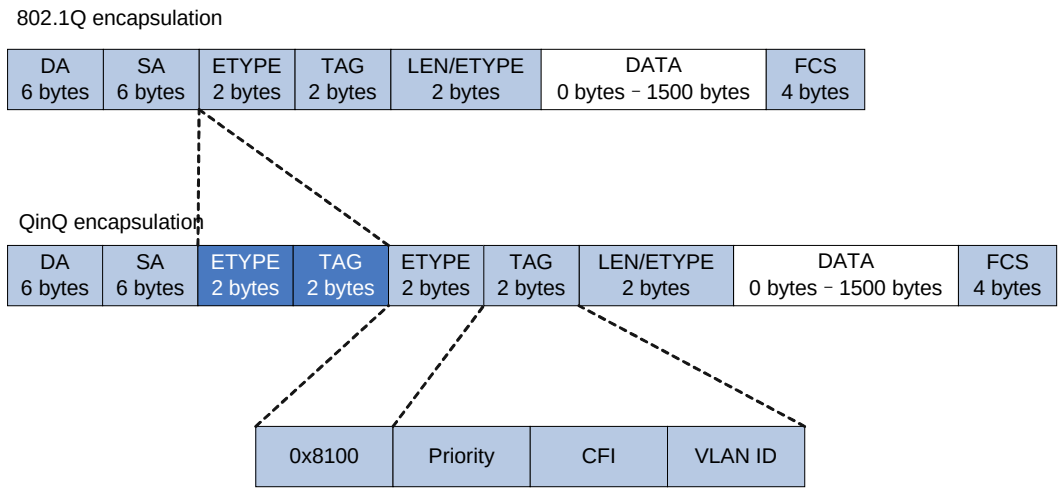
```
huawei(config)#acl 3001
huawei(config-acl-basic-3001)rule 1 deny icmp destination 10.10.10.101 0
huawei(config-acl-basic-3001)rule 2 deny tcp destination 10.10.10.101 0
destination-port eq telnet
huawei(config-acl-basic-3001)quit
huawei(config)#packet-filter inbound ip-group 3001 rule 1 port 0/2/0
huawei(config)#packet-filter inbound ip-group 3001 rule 2 port 0/2/0
huawei(config)#save
```

To create a link-layer ACL rule that allows data packets with protocol type 0x8863 (pppoe-control message), VLAN ID 12, CoS 1, source MAC address 2222-2222-2222, and destination MAC address 00e0-fc11-4141 to pass, do as follows:

```
huawei(config)#acl 4001
huawei(config-acl-link-4001)rule 1 permit type 0x8863 cos 1 source 12
2222-2222-2222 0000-0000-0000 destination 00e0-fc11-4141 0000-0000-0000
huawei(config-acl-basic-4001)quit
huawei(config)#save
```

Assuming that the packet sent from port 0/2/0 to the MA5600T/MA5603T/MA5608T is the QinQ packet containing two VLAN tags, to change the CoS priority in the outer VLAN tag (VLAN ID: 10) to 5, do as follows:

Figure 14-42 QinQ packet format




```
huawei(config)#acl 5001
huawei(config-acl-user-5001)#rule 1 permit 8100 ffff 16
```

NOTE

The type value of a QinQ packet varies with vendors. Huawei adopts the default 0x8100. As shown in Figure 14-42, the offset of this type value should be 16 bytes.

```
huawei(config-acl-user-5001)#rule 10 permit 0a ff 19
huawei(config-acl-user-5001)#quit
```

NOTE

"19" indicates the ADN operation after an offset of 19 bytes with the header of the packet as the base. "0a" refers to the value of the inner tag field of the QinQ packet. In this example, the second byte of the inner tag field is a part of the VLAN ID, which is exactly the value of the inner VLAN ID (VLAN 10).

```
huawei(config)#traffic-priority inbound user-group 5001 cos 5 port 0/2/0
```

Configuring ACL Matching for PPPoE Packets

This topic describes the format of PPPoE packets and how to match user-defined ACLs for various PPPoE packets.

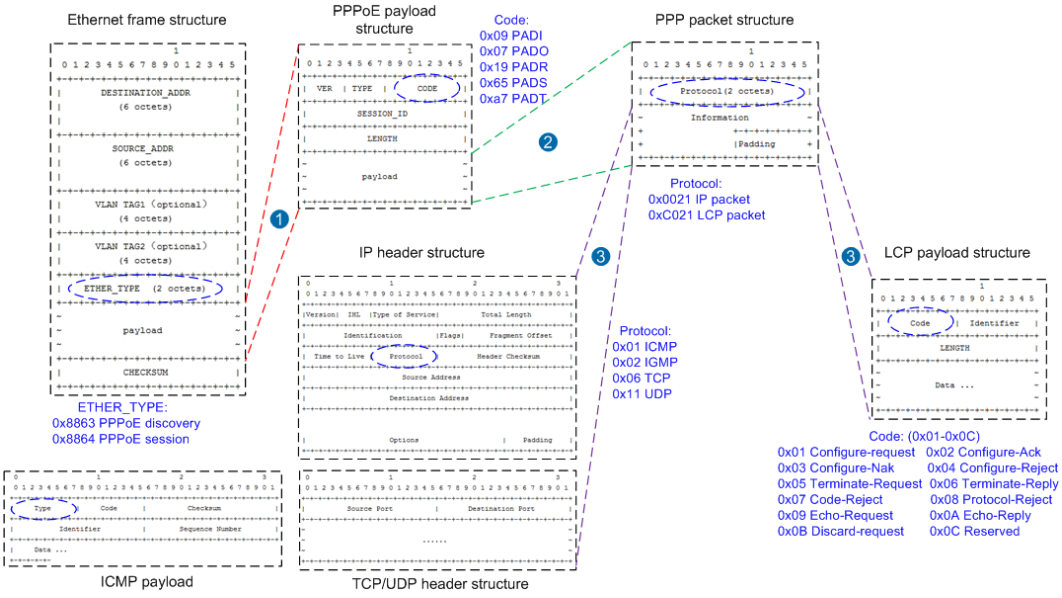
Prerequisites

You are familiar with the 14.10.5 Matching Principle for the User-defined ACL Rule.

Context

Figure 14-43 shows the PPPoE packet format and offset value of each field.

Figure 14-43 PPPoE packet format and offset value of each field



Basic knowledge about PPPoE packets:

- In PPPoE-based data transmission, data packets are carried in the payload of ETH packets for transmission.

- The whole PPPoE-based transmission process has two stages: discovery (ETHER_TYPE = 0x8863) and session (ETHER_TYPE = 0x8864).
- In the discovery stage, there are four types of packets: PADI (code = 0x09), PADO (code = 0x07), PADR (code = 0x19), and PADS (code = 0x65).
- In the session stage, data is transmitted through PPP and PPP data is carried in the payload of PPPoE packets.
- PPP data frames can carry IP data packets (protocol = 0x0021) or LCP data packets (protocol = 0xC021).
- Based on the value of the protocol field in the IP header of IP data packets, the data transmission mode can be determined. The data transmission modes include:
  - 0x01 ICMP
  - 0x02 IGMP
  - 0x06 TCP
  - 0x11 UDP
- Based on the value (value range: 0x01-0x0C) of the code field of the LCP data packets, the packet encapsulation type can be determined.

Configuration Guideline

1. Determine the stage status of PPPoE packets, discovery or session.
2. Determine the protocol used to carry PPP packets, IP or LCP.
3. Determine the transmission mode (UDP, IGMP, or others) or encapsulation type of data packets.
4. Determine the offset value of each field.

Procedure

- **Match ACLs for PADI packets with a specified source MAC address and a single VLAN tag.**

  As shown in Figure 14-43, PADI packets contain the following information:
  - Packets are in the discovery stage.
  - Since the value of the ETHER_TYPE field of ETH packets is 0x8863, and packets carry only one single VLAN tag, the offset value is 16 (6 + 6 + 4 = 16).
  - Since the value of the code field of PPPoE packets is 0x09, the offset value is 19 (16 + 2 + 1 = 19).

  Based on the 14.10.5 Matching Principle for the User-defined ACL Rule, the corresponding ACLs can be determined.

  Example: Run the following commands to match ACLs for PADI packets with the source MAC address of 0x0011-2233-4455 and a single VLAN tag:

```
huawei(config)#acl 5001
huawei((config-acl-user-5001))#rule permit 00112233 FFFFFFFF 6 4455 FFFF 10 8863 FFFF
16 09 FF 19 non-ipoe
```

- **Match ACLs for PPPoE echo-request packets with two VLAN tags.**

  As shown in Figure 14-43, PPPoE echo-request packets contain the following information:
  - Packets are in the session stage and LCP is used for packet transmission.
  - Since the value of the ETHER_TYPE field of ETH packets is 0x8864, and packets carry two VLAN tags, the offset value is 20 (6 + 6 + 4 + 4 = 20).

- Since the value of the protocol field of PPP packets is 0xC021, the offset value is 28 (20 + 2 + 6 = 28).
- Since the value of the code field of LCP packets is 0x09, the offset value is 30 (28 + 2 = 30).

Based on the 14.10.5 Matching Principle for the User-defined ACL Rule, the corresponding ACLs can be determined.

Example: Run the following commands to match ACLs for PPPoE echo-request packets with two VLAN tags:

```
huawei(config)#acl 5001
huawei((config-acl-user-5001))#rule permit 8864 FFFF 20 C021 FFFF 28 09 FF 30 non-ipoe
```

- **Match ACLs for packets with a specified inner VLAN ID and a specified UDP source port number.**

  As shown in Figure 14-43, the packets contain the following information:

  - Packets are in the session stage and IP UDP is used for packet transmission.
  - Since the value of the ETHER_TYPE field of ETH packets is 0x8864, and packets carry two VLAN tags, the offset value is 20 (6 + 6 + 4 + 4 = 20).
  - Since the value of the protocol field of PPP packets is 0x0021, the offset value is 28 (20 + 2 + 6 = 28).
  - Since the value of the protocol field in the IP header is 0x11, the offset value is 39 (28 + 2 + 9 = 39).
  - Based on the UDP source port number, the offset value is 50 (39 + 1 + 10 = 50).

  Based on the 14.10.5 Matching Principle for the User-defined ACL Rule, the corresponding ACLs can be determined.

  Example: Run the following commands to match ACLs for PPPoE packets whose inner VLAN ID is 291 (0x123) and UDP source port number is 94 (0x5e):

```
huawei(config)#acl 5001
huawei((config-acl-user-5001))#rule permit 0123 0FFF 16 8864 FFFF 20 0021 FFFF 28 11
FF 39 5E FF 50 non-ipoe
```

- **Match ACLs for PPPoE-based PING (echo-request) packets with specified inner and outer VLAN IDs.**

  As shown in Figure 14-43, the packets contain the following information:

  - Packets are in the session stage and IP ICMP is used for packet transmission.
  - Since the value of the ETHER_TYPE field of ETH packets is 0x8864, and packets carry two VLAN tags, the offset value is 20 (6 + 6 + 4 + 4 = 20).
  - Since the value of the protocol field of PPP packets is 0x0021, the offset value is 28 (20 + 2 + 6 = 28).
  - Since the value of the protocol field in the IP header is 0x11, the offset value is 39 (28 + 2 + 9 = 39).
  - Since the value of the type field in the ICMP header is 0x08, the offset value is 50 (39 + 1 + 10 = 50).

  Based on the 14.10.5 Matching Principle for the User-defined ACL Rule, the corresponding ACLs can be determined.

  Example: Run the following commands to match ACLs for PPPoE-based PING (echo-request) packets whose outer and inner VLAN IDs are 256 (0x100) and 291 (0x123), respectively:

```
huawei(config)#acl 5001
huawei((config-acl-user-5001))#rule permit 0100 0FFF 12 0123 0FFF 16 8864 FFFF 20 0021 FFFF 28 01 FF 39 08 FF 50 non-ipoe
```

----End

Controlling the Traffic Matching an ACL Rule

This topic describes how to control the traffic matching an ACL rule on a specified port, and process the traffic that exceeds the limit, such as adding the DSCP tag or dropping the packet directly.

Prerequisite

The Configuring an ACL Rule, and the port for traffic limit is working in the normal state.

Context

- The traffic statistics are only effective for the permit rules of an ACL.
- The limited traffic must be an integer multiple of 64 kbit/s.

Procedure

Run the **traffic-limit** command to control the traffic matching an ACL rule on a specified port.

Use the **target-rate** parameter to set the fixed maximum rate of the port, or use CAR parameters to set a rate for trTCM-based ports. The two rates cannot be set at a time. Run this command to set the action to be taken when the traffic received on the port exceeds the limited value. Two options are available:

- **drop**: Drop the traffic that exceeds the limited value.
- **remark-dscp** *value*: To set the DSCP priority for the traffic that exceeds the limited value, use this parameter.

Step 1 Run the **display qos-info traffic-limit port** command to query the traffic limit information on the specified port.

----End

Example

To limit the traffic that matches ACL 2001 received on port 0/2/0 to 512 kbit/s, and add the DSCP priority tag (af1) to packets that exceed the limit, do as follows:

```
huawei(config)#traffic-limit inbound ip-group 2001 512 exceed remark-dscp af1 port 0/2/0
//"af1" represents a dscp type: Assured Forwarding 1 service (10).
huawei(config)#display qos-info traffic-limit port 0/2/0
traffic-limit:
port 0/2/0:
 Inbound:
   Matches: Acl 2001 rule 5    running
     Target rate: 512 Kbps
     Exceed action: remark-dscp af1
```

Adding a Priority Tag to the Traffic Matching an ACL Rule

This topic describes how to add a priority tag to the traffic matching an ACL rule on a specified port so that the traffic can obtain the service that matches the specified priority. The priority tag type can be ToS, DSCP, or 802.1p.

Prerequisite

The Configuring an ACL Rule, and the port for traffic limit is working in the normal state.

Context

- The traffic statistics are only valid to permit rules of an ACL.
- The ToS and the DSCP priorities are mutually exclusive. Therefore, they cannot be configured at the same time.

Procedure

Run the **traffic-priority** command to add a priority tag to the traffic matching an ACL rule on a specified port.

Step 1 Run the **display qos-info traffic-priority port** command to query the configured priority.

----End

Example

To add a priority tag to the traffic that matches ACL 2001 received on port 0/2/1, and the DSCP priority and local priority of the traffic are 10 (af1) and 0 respectively, do as follows:

```
huawei(config)#traffic-priority inbound ip-group 2001 dscp af1 local-precedence 0 port 0/2/1
huawei(config)#display qos-info traffic-priority port 0/2/1

traffic-priority:
port 0/2/1:
 Inbound:
   Matches: Acl 2001 rule 5 running
     Priority action: dscp af1 local-precedence 0
```

Enabling the Statistics Collection of the Traffic Matching an ACL Rule

This topic describes how to enable the statistics collection of the traffic matching an ACL rule, analyzing and monitoring the traffic.

Prerequisite

The Configuring an ACL Rule, and the port for traffic statistics is working in the normal state.

Context

The traffic statistics are only valid to permit rules of an ACL.

Procedure

Run the **traffic-statistic** command to enable the statistics collection of the traffic matching an ACL rule on a specified port.

Step 1 Run the **display qos-info traffic-mirror port** command to query the statistics information about the traffic matching an ACL rule on a specified port.

----End

Example

To enable the statistics collection of the traffic that matches ACL 2001 received on port 0/19/0, do as follows:

```
huawei(config)#traffic-statistic inbound ip-group 2001 port 0/19/0
huawei(config)#display qos-info traffic-statistic port 0/19/0

traffic-statistic:
port 0/19/0:
 Inbound:
   Matches: Acl 2001 rule 5    running
     0 packet
```

Enabling the Mirroring of the Traffic Matching an ACL Rule

This topic describes how to mirror the traffic matching an ACL rule on a port to a specified port. Mirroring does not affect packet receipt and transmission on the mirroring source port. You can monitor the traffic of the mirroring source port by analyzing the traffic that passes the mirroring destination port.

Prerequisite

The Configuring an ACL Rule, and the port for traffic mirroring is working in the normal state.

Context

- The traffic statistics are only valid to permit rules of an ACL.
- The destination mirroring port cannot be an aggregation port.
- The system supports only one mirroring destination port and the mirroring destination port must be the upstream port.

Procedure

Run the **traffic-mirror** command to enable the mirroring of the traffic matching an ACL rule on a specified port.

Step 1 Run the **display qos-info traffic-mirror port** command to query the mirroring information about the traffic matching an ACL rule on a specified port.

----End

Example

To mirror the traffic that matches ACL 2001 received on port 0/2/1 to port 0/19/0, do as follows:

```
huawei(config)#traffic-mirror inbound ip-group 2001 port 0/2/1 to port 0/19/0
huawei(config)#display qos-info traffic-mirror port 0/2/1

traffic-mirror:
port 0/2/1:
 Inbound:
   Matches: Acl 2001 rule 5    running
   Mirror to: port 0/19/0
```

Enabling the Redirection of the Traffic Matching an ACL Rule

This topic describes how to redirect the traffic matching an ACL rule on a specified port. After this operation is executed successfully, the original port does not forward the traffic matching the ACL rule, but the specified port forwards the traffic.

Prerequisites

The Configuring an ACL Rule, and the port for redirection is working in the normal state.

Context

- The traffic statistics are only valid to permit rules of an ACL.
- Currently, the service ports support only redirection of the traffic matching the ACL rule to upstream ports. The upstream ports support only redirection of the traffic matching the ACL rule to ports on the board of the same type.

Procedure

Run the **traffic-redirect** command to redirect the traffic matching an ACL rule on a specified port.

Step 1 Run the **display qos-info traffic-redirect port** command to query the redirection information about the traffic matching an ACL rule on a specified port.

----End

Example

To redirect the traffic that matches ACL 2001 received on port 0/19/0 to port 0/19/1, do as follows:

```
huawei(config)#traffic-redirect inbound ip-group 2001 port 0/19/0 to port 0/19/1
huawei(config)#display qos-info traffic-redirect port 0/19/0
traffic-redirect:
port 0/19/0:
 Inbound:
   Matches: Acl 2001 rule 5    running
    Redirected to: port 0/19/1
```

14.11 ACLv6

This topic describes the aspects unique to access control list for IP version 6 (ACLv6), and the differences between ACLv6 and ACLv4. For details about the specifications and principles of ACLv6, see 14.10 .

14.11.1 Comparison Between ACLv6 and ACLv4

On the MA5600T/MA5603T/MA5608T, the application and the configuration processes of ACLv6 are the same as those of ACLv4, except the following differences regarding specifications and commands:

- ACLv4 supports segmented packets, while ACLv6 does not.
- IPv6 and IPv4 have different IP address formats and packet formats, so the **ipv6** parameter must be specified when you are configuring basic ACLs and advanced ACLs for IPv6. When the **ipv6** parameter is specified, the system supports both ACLv4 and ACLv6.
- Regarding user-defined ACLs, packets are matched based on the packet type defined by users. Because IPv4 and IPv6 have the same command format for specifying the packet type in user-defined ACLs, configurations do not differentiate between IPv6 user-defined ACLs and IPv4 user-defined ACLs. When user-defined ACLs are used for filtering packets, however, the protocol type of the packets must be matched against the protocol type specified in the ACL rules.

14.12 HQoS

Traditional quality of service (QoS) schedules traffic based on port. Hierarchical quality of service (HQoS) not only controls port traffic at finer service granularities but also schedules traffic based on service priorities.

14.12.1 Overview

Definition

Hierarchical quality of service (HQoS) is a technology that uses a multi-level scheduling mechanism to guarantee the bandwidth of various services for multiple users. It usually applies to Open Access networks.

The following HQoS scheduling models are available:

- The first level of HQoS guarantees the bandwidth of services for a user, and the second level guarantees the bandwidth of all users that use the same types of services.
- The first level of HQoS guarantees the bandwidth of services for a user, and the second level guarantees the bandwidth of all services for the user.

Background

In most scenarios, traditional QoS identifies service types on a network and provides services at different levels. As users and service types increase, traditional QoS faces the following challenges:

- Traditional QoS schedules traffic based on port bandwidth. Traffic management is therefore sensitive to the service level rather than users, which applies to traffic at the network core side rather than traffic at the service access side.
- Traditional QoS is unable to uniformly manage or hierarchically schedule various services or multiple users.

Purpose

HQoS controls traffic in a user-specific and service-specific manner, and therefore guarantees bandwidth of various services for multiple users. In addition, it provides the committed information rate (CIR) and peak information rate (PIR) of various services for each user.

Benefits

Benefits to carriers

Unlike traditional QoS which schedules traffic based on a port, HQoS implements QoS on a port at finer granularities (users and service flows). Therefore, HQoS enables a carrier to guarantee QoS for enterprises and contracted users, provide guaranteed bandwidths and service packages for more users, and achieve higher profitability.

Benefits to users

HQoS ensures that the bandwidth designated for a user is not affected by other users.

14.12.2 Open Access

Hierarchical quality of service (HQoS) mainly applies to Open Access networks. This topic describes the Open Access network model for better understanding of HQoS.

Definition and Benefits

Open Access provides a network business model that separates the physical bearer network from the service network. The infrastructure of an Open Access network, including passive infrastructure (optical fibers, equipment room premises, and cables) and active network devices, is built by a nation or an operator authorized by the nation. Retail service providers (RSPs) directly lease bandwidth on the infrastructure network to provide service packages to end-users.

In the traditional model, an operator builds and operates its own network and delivers services to end-users. Unlike the traditional model, Open Access builds a layered network over which separate RSPs deliver their services.

Open Access brings the following benefits:

- Maximizes the freedom of choice for end users. End-users have more services to choose from and can even switch from one service provider to another without changing their home terminals (such as their ONTs).
- Lowers investment risks for RSPs. The business model of Open Access greatly shortens the cycle of return on investment (ROI). The traditional business model of operators, who usually have monopoly over their networks, requires an ROI cycle of 8-10 years. The Open Access business model shortens the ROI cycle to 1-2 years. Hence Open Access lowers the investment entrance level and risks for RSPs, and promotes competition and innovation.

- Opens up a wider arena for RSPs. RSPs no longer need to build the infrastructure network and are able to focus on innovation and competition of services and contents.

Open Access Modes

In a broad sense, there are two Open Access modes: the physical open access mode (LLU) and the bit stream open access mode.

- LLU Access Mode: This is a layer-one physical open access mode. In this mode, RSPs lease duct resources such as copper loops and optical fibers. One line cannot carry the services of multiple RSPs. To subscribe to services of different content providers (CPs), a user needs to apply for respective lines.
- Bit Stream Access Mode: In the bit stream access mode, RSPs are separated from the infrastructure network. RSPs purchase bandwidth on the infrastructure network and provide service packages to end-users.

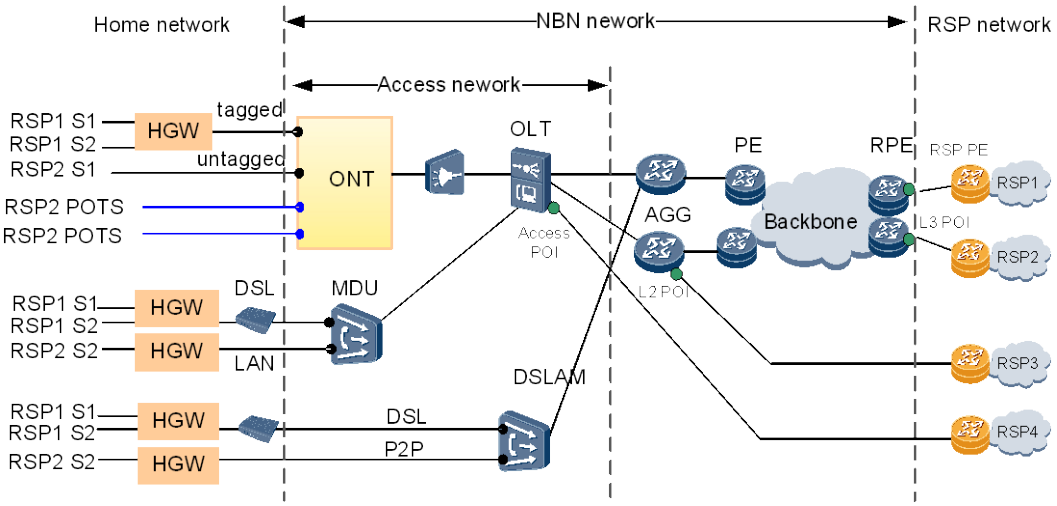
Wholesale is a special bit stream access mode. RSPs lease physical terminals, and one terminal belongs to one RSP. Wholesale allows for simpler terminal management.

The bit stream access mode presents multiple requirements and challenges to FTTx devices. Therefore, the following part of this document will focus on the bit stream access mode.

End-to-end Open Access network

Figure 14-44 shows an end-to-end Open Access network (Bit Stream Access Mode).

Figure 14-44 End-to-end Open Access network



NBN: short for national broadband network. An NBN is usually constructed using the Open Access network model.

POI: short for point of interconnection. It is a point where a carrier network and retail service provider (RSP) network interconnects. Based on network layers, there are access POI, Layer 2 POI, and Layer 3 POI.

An end-to-end Open Access network consists of the RSP networks at both ends and the NBN open channel network in between. The NBN open channel network includes the ONT, OLT,

and aggregation devices such as the user-end provider edge (UPE) and access aggregation gateway (AGG).

An RSP needs to purchase physical ports and logical bandwidth to access the NBN open channel network. The physical ports are the openings at both ends of the open channel.

- User network interfaces (UNIs). They are usually the ports on user-side devices, such as the ETH/POTS/Wi-Fi ports on an ONT, xDSL/LAN ports on a multi-dwelling unit (MDU), and xDSL/P2P ports on a digital subscriber line access multiplexer (DSLAM). UNIs connect upper-layer devices to user terminals or RSPs' home gateways.
- External network-to-network interfaces (ENNIs), also called point of interconnection (POI), are the interconnection points between an operator network and an RSP network. POIs can be the upstream GE or 10GE ports on an OLT, aggregation device, or backbone network device.

Because RSPs will share one physical network, the RSPs purchase logical bandwidth on the network and use the bandwidth as service channels to provision services to end-users. The logical bandwidth includes the bandwidth on the UNIs for access users and the aggregation bandwidth on the ENNIs. The logical bandwidth is expressed in the unit of bit/s.

Openness and fairness are two key points of a bit stream Open Access network. Openness means that an ONT on user side can be shared by multiple RSPs for provisioning services. Also, on network side the ENNIs can provide various openings to meet different network requirements. Fairness means that users of the same RSP can be isolated from each other, and RSPs can be isolated from each other, through HQoS.

14.12.3 Basic Concepts

Concept	Description
HQoS user	A basic unit of bandwidth guarantee and traffic scheduling. Alternatively, a type of user service can also be a basic unit.
HQoS user group	A group of HQoS users. The committed information rate (CIR) of HQoS users is guaranteed by scheduling an HQoS user group.
Retail service provider (RSP)	Unlike carriers, RSPs focus on service content rather than infrastructure network construction. They lease bandwidth to quickly provision user services.
Open Access	A network construction mode that separates the physical bearer network from the service network. RSPs directly lease bandwidth on a infrastructure network where Open Access has been deployed to quickly provide service packages for customers.
CAR group	A combination of service flows for unified QoS control, with configurable parameters such as CIR and peak information rate (PIR). Generally, a CAR group is specified for limiting the rate of triple-play services of a specific household user.

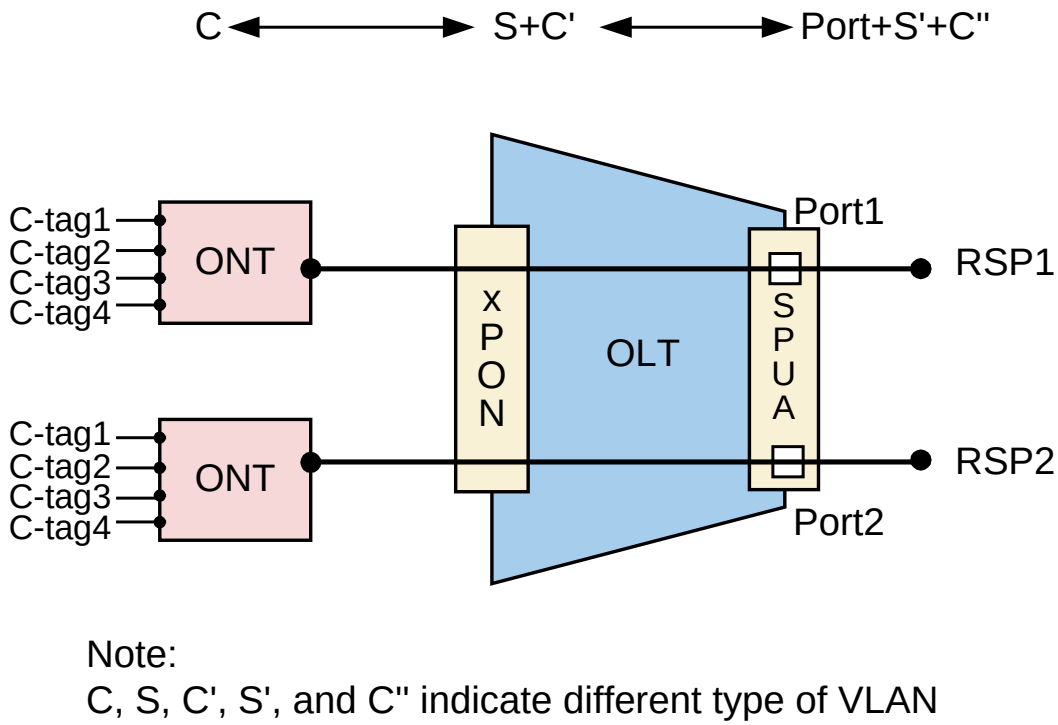
14.12.4 HQoS Service Model (Based on Port+VLAN)

Figure 14-45 shows an HQoS service processing model based on Port+VLAN. A Layer 2 leased line is used, between the ingress port on the optical network terminal (ONT) and the

uplink port of the optical line terminal (OLT). On the line, services are forwarded based on S-VLAN+C-VLAN tags.

- The ONT attaches C-VLAN tags to services based on a port or service type to identify users.
- After receiving the services, the OLT translates the C-VLAN tags into C-VLAN' tags based on PON port information and adds S-VLAN tags to the services based on service types. The services are identified based on their S-VLAN+C-VLAN' tags.
- After services are forwarded to a port on the SPUA board that is interconnected to an RSP network, the SPUA board translates the S-VLAN+C-VLAN' tags into S-VLAN' +C-VLAN" tags based on the RSP planning requirements configured on the port.

Figure 14-45 HQoS service model based on Port+VLAN



The Layer 2 leased line can be considered as designated to an HQoS user with the following QoS requirements:

1. CIR can be ensured when network congestion occurs, and PIR can be ensured when the network is idle.
2. The services of the user occupy bandwidth based on service priorities. Specifically, services with a higher priority can preferentially occupy CIR and PIR resources.

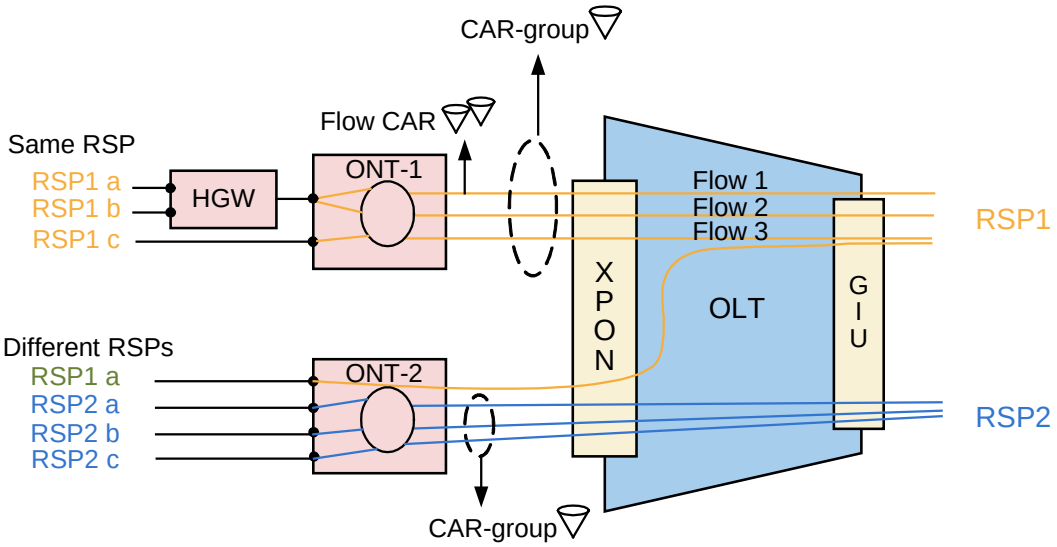
14.12.5 HQoS Service Model (Based on a CAR Group)

Figure 14-46 shows a hierarchical quality of service (HQoS) service processing model based on a committed access rate (CAR) group.

- Installed at a user's home, an optical network terminal (ONT) can run services provided by one retail service provider or RSP (such as OTN-1 shown in the figure), or services provided by multiple RSPs (such as OTN-2 shown in the figure).

- On the ONT, triple play services are provisioned and each type of service is transmitted through one GEM port. On the optical line terminal (OLT), each type of service is mapped into one service flow.
- There is no limit on the VLAN tag translation mode. An S-VLAN+C-VLAN tagging mode, however, is recommended (the S-VLAN tag indicates an RSP while the C-VLAN tag indicates a service type).

Figure 14-46 HQoS service model based on a CAR group



Notes:
The a, b, and c indicate different service types.

This HQoS service model can address the following QoS requirements:

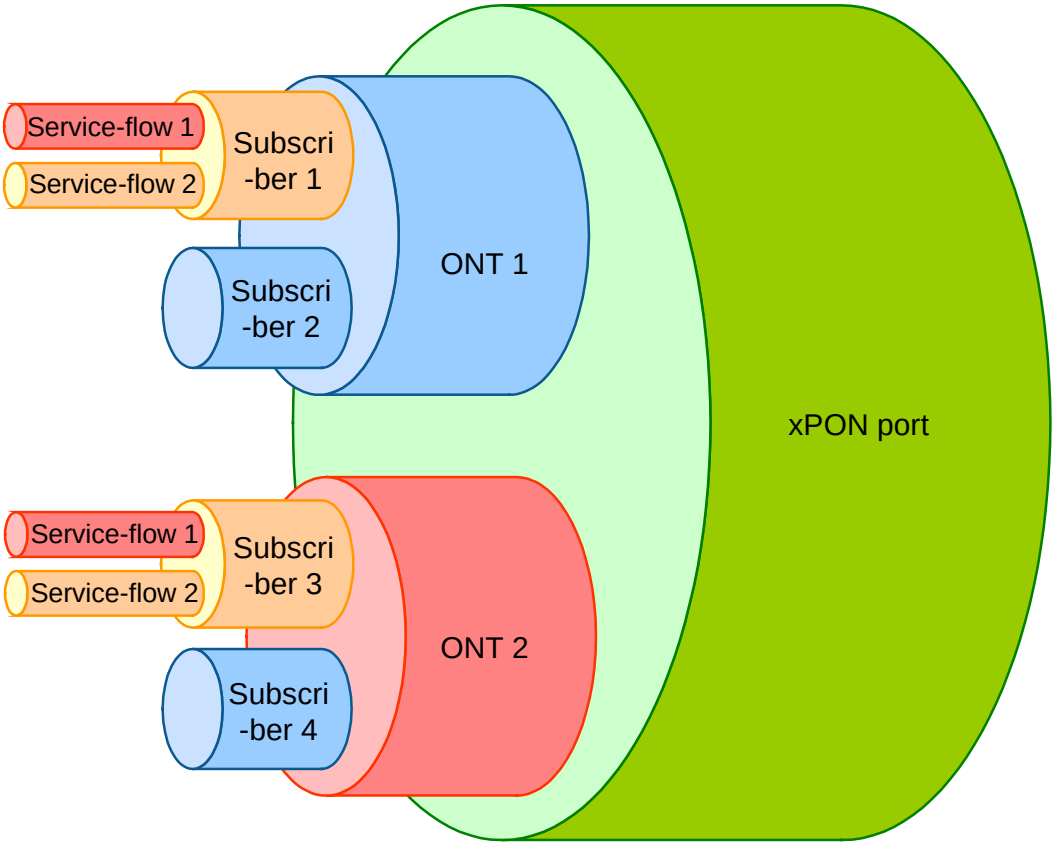
1. For each type of service, CIR can be ensured when network congestion occurs, and PIR can be ensured when the network is idle.
2. When services are provided by only one RSP, the total user bandwidth can be limited.
3. When services are provided by multiple RSPs, the total lease bandwidth of each RSP can be limited.

14.12.6 HQoS Service Model (xPON Board)

Figure 14-47 shows the HQoS service model of an xPON board.

- One xPON port connects to multiple ONTs.
- One ONT connects to multiple users.
- One user is provisioned with multiple services.

Figure 14-47 HQoS service model (xPON board)



As shown in the figure, this model supports the specifications for HQoS:

- Rate limitation on every user and every service
- Bandwidth scheduling between multiple services of every user
- Rate limitation on every user
- Bandwidth scheduling on every ONT
- Rate limitation on every ONT
- Scheduling between multiple ONTs connected to every xPON port

14.12.7 Implementation Principle

Implementation Principle of HQoS based on Port+VLAN

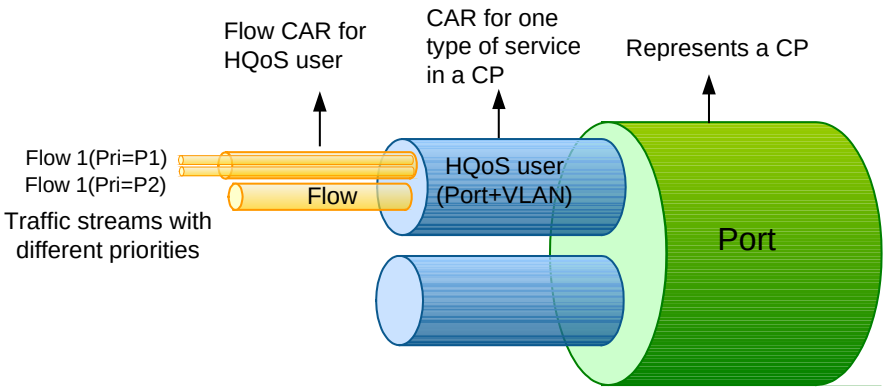
Implementation Model of HQoS based on Port+VLAN

Figure 14-48 shows an implementation model of HQoS based on Port+VLAN. In the model, each HQoS user is mapped to one Layer 2 service flow.

The core of the implementation principle of HQoS based on Port+VLAN is a two-level service rate limitation mechanism on the SPUA board. The following describes the mechanism:

1. The committed access rate (CAR) is limited for an HQoS user, and the service packets are marked with colors based on committed information rate (CIR) and peak information rate (PIR). If priority-based CAR is enabled (with CAR thresholds configured), packets with a higher priority will be marked green and transmitted.
2. Color-based CAR is performed on an HQoS user group (a type of service provided by an RSP and identified by port+VLAN tags). In this manner, the bandwidth for a type of service can be ensured among the total leased RSP bandwidth, HQoS users can be isolated from each other, the CIR can be ensured for the services of HQoS users, and the PIR can also be ensured if bandwidth resources are sufficient.

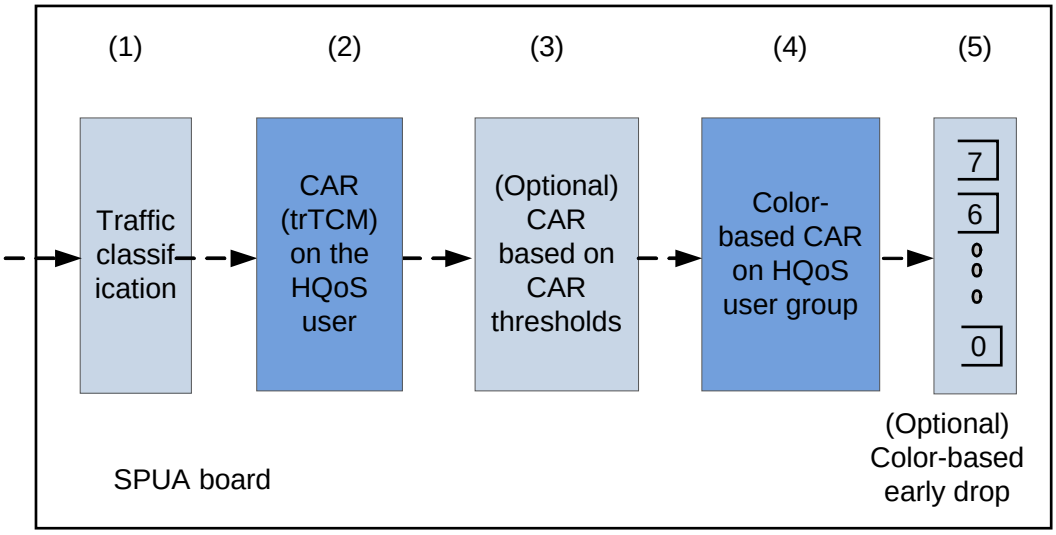
Figure 14-48 Implementation model of HQoS based on Port+VLAN



Implementation Solution

Figure 14-49 shows the implementation solution of HQoS based on Port+VLAN.

Figure 14-49 Procedure for implementing HQoS based on Port+VLAN



The following describes how HQoS is implemented using the SPUA board:

1. The SPUA board performs traffic classification on received packets.

On an optical line terminal (OLT), each HQoS user is mapped to one Layer 2 service flow. When a service packet from the uplink direction enters the OLT, the SPUA board on the OLT performs traffic classification on it based on related packet parameters and determines the HQoS user to which the service packet belongs.

- On the network-to-network interface (NNI) side (or an uplink port of the SPUA board), traffic classification can be performed based on port+S-VLAN+C-VLAN' tags.
- On the user-to-network interface (UNI) side, traffic classification can be performed based on S-VLAN+C-VLAN' tags.

For detailed principles of traffic classification, see 14.5.2 Implementation Principle.

2. The SPUA board performs CAR on the HQoS user (flow CAR) using a two rate three color marker (trTCM) algorithm.

   The traffic profile bound to the Layer 2 service flow (that is, the HQoS user) is used for CAR. Specifically, a packet is marked with a color based on the CIR and PIR.

   - If the packet rate is lower than or equal to the CIR, the SPUA board will mark the packet green (and will transmit it).
   - If the packet rate is higher than the CIR, but is lower than or equal to the PIR, the SPUA board will mark the packet yellow (and will transmit it).
   - If the packet rate is higher than the PIR, the SPUA board will directly drop the packet.

   For detailed principles of the trTCM algorithm, see Dual-Token Bucket (trTCM) Principle.

3. (Optional) The SPUA board performs flow CAR based on CAR thresholds, and marks user packets with colors based on packet priorities (configurable through the **car-threshold** command).

   The SPUA board marks the packets of the HQoS user with colors based on the CIR and PIR settings. (If an enhanced trTCM algorithm is used, thresholds for 4 priorities will be supported.) A packet will be marked green if the packet rate is lower than the CIR, or yellow if the packet rate is between the CIR and the PIR, or will be directly dropped if the packet rate is higher than the PIR. When marking packets with colors, the SPUA board also differentiates packet priorities. Packets with a higher priority can preferentially use the CIR and PIR bandwidth specified for the HQoS user.

   For detailed principles of packet color marking based on CAR thresholds, see Enhanced trTCM: Marking Packet with Color Based on CAR Thresholds.

4. The SPUA board performs CAR on HQoS user groups (group CAR) based on packet colors (configurable through the **car-port** ***portid*** **vlan** command).

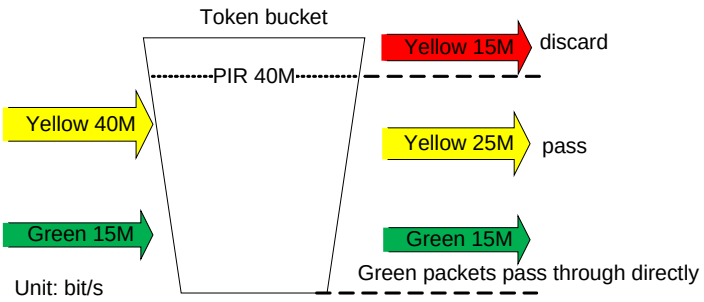
   As shown in Figure 14-50, CAR is performed on yellow packets and green packets that have been identified in step 2. Specifically,

   a. Green packets will be transmitted.

   b. All yellow packets can be transmitted if their rate is within the remaining bandwidth (Remaining bandwidth = PIR - Rate of green packets).

   c. If the rate of yellow packets exceeds the remaining bandwidth, some yellow packets can be transmitted using the remaining bandwidth while the excessive yellow packets will be dropped.

Figure 14-50 CAR on yellow packets and green packets



5. (Optional) The SPUA board performs color-based early drop on an egress port.

   On the egress port, the SPUA board puts all HQoS users into the same priority queue. The queue deploys a strict priority-based scheduling policy, and color-based early drop will be performed based on the queue. When the rate of yellow packets reaches the specific drop threshold, yellow packets will be dropped to ensure low latency of green packets. Note that the drop threshold is 50% for yellow packets and 100% for green packets. The two drop thresholds cannot be modified manually.

   To put HQoS users into one priority queue, the following configuration methods can be used:

   - In fiber to the building (FTTB) scenarios where both S-VLAN and double C-VLAN tags are attached to packets, the outer S-VLAN tags of HQoS users can be set to a specific priority so that all HQoS users can be in the same queue based on the priority.
   - In fiber to home (FTTH) scenarios where packets are single-tagged, all priorities of single tags can be mapped into one port queue.

   For details on color-based early drop, see Color-based Early Drop.

Implementation Principle of CAR-Group-based HQoS

Implementation Model of CAR-Group-based HQoS

A CAR group is a combination of traffic streams for unified QoS control, with configurable parameters such as CIR and PIR. One of its typical applications is for multiple services (IPTV, Internet access, and voice services) of home users. Using the CAR group, QoS based on the home user instead of based on each service is implemented.

Figure 14-51 illustrates the model of CAR-group-based HQoS supported by the MA5600T/MA5603T/MA5608T.

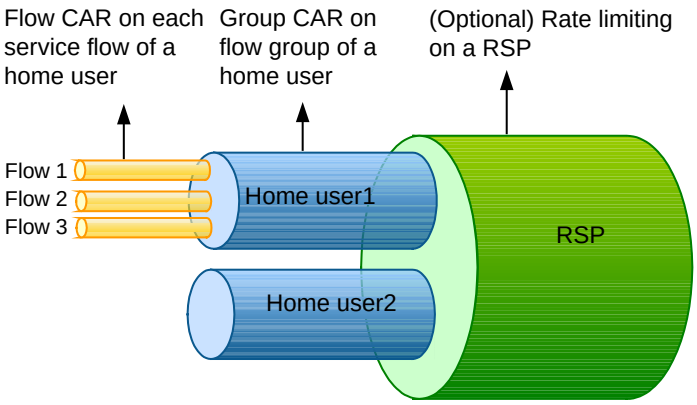
HQoS based on a CAR group implements two-level CAR for services:

1. The first level performs flow CAR on various service flows of a home user.
2. The second level performs group CAR on flow groups of home users.

If required, a level of CAR can be performed specifically for retail service providers (RSPs).

For example, the two-level CAR can achieve the following settings: limiting a user's Internet access rate, multicast service rate, and voice service rate to 2 Mbit/s, 4 Mbit/s, and 128 kbit/s respectively, and at the same time limiting the total bandwidth of the user to 5 Mbit/s.

Figure 14-51 Implementation model of CAR-group-based HQoS

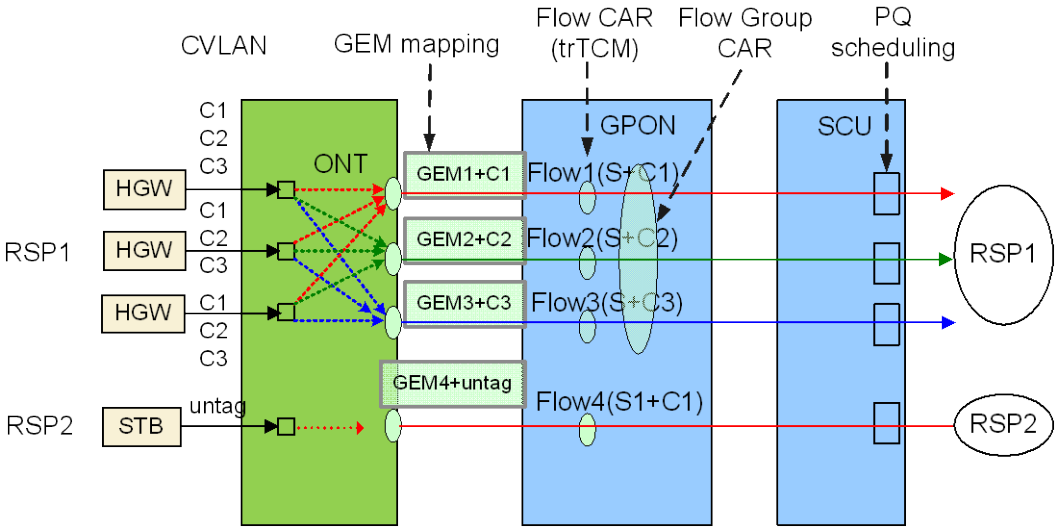


Implementation Solution

Figure 14-52 shows how HQoS is implemented based on a CAR group in the uplink direction.

In the downlink direction, a reverse process is performed. This topic only describes the implementation process in the uplink direction.

Figure 14-52 Flowchart for implementing HQoS based on a CAR group (in the uplink direction)



RSP: retail service provider

HGW: home gateway

STB: set-top box

GEM: G-PON encapsulation method

trTCM: two rate three color marker

Table 14-16 Implementation principle of HQoS based on a CAR group in the uplink direction

Subject	Operation	Result
HGW	Identifies different types of services provided by the same RSP based on C-VLAN tags. In Figure 14-52, C1, C2, and C3 indicate different service types.	It receives various types of services and identifies them based on the C-VLAN tags of service packets.
ONT	Maps each type of service into one GEM port based on VLAN tags. In the figure, services tagged as C1 are mapped to GEM port 1 while services tagged as C2 are mapped to GEM port 2.	It identifies various service packets based on GEM port+C-VLAN tags.
	Binds each GEM port to one T-CONT. Performs priority queuing (PQ) on T-CONTs. (To prevent queuing differences, the class of service, or CoS, must have been uniformly configured.)	Guarantees the uplink bandwidth of each type of service on a T-CONT.
Service board of the OLT	Performs traffic classification and priority processing. Each GEM port+C-VLAN tag identifies one service flow. The priority of the service flow can be configured (if the user-side priority is not trusted), or directly duplicated from the user-side priority (if the user-side priority is trusted). Performs two-level CAR: 1. Flow CAR (rate limiting on each service flow) using the trTCM algorithm. For details, see 14.7.4 Traffic Policing Mode. 2. Group CAR (rate limiting on all service flow groups provided by the same RSP) based on the packet color.	It identifies service flows through traffic classification, preparing for CAR. Flow CAR is based on the packet color that is marked based on the CIR and PIR. A packet is marked green (and will be transmitted) if the packet rate is lower than the CIR, is marked yellow (and will be transmitted) if the packet rate is higher than the CIR but is lower than or equal to the PIR, or is dropped if the packet rate is higher than the PIR. Group CAR is implemented using a single-token leaky bucket mechanism. With the single-token leaky bucket mechanism, drop thresholds can be set for different priorities of yellow packets to ensure that yellow packets with a higher priority can be preferentially transmitted. (The drop thresholds are the half of the CAR thresholds set using the **CAR-threshold** command). **NOTE** The PIR of a flow group must be higher than or equal to the sum of the CIRs of all flows. The two-level CAR mechanism guarantees the QoS for each type of

Subject	Operation	Result
		service and for all services of a home user.
Uplink port of the OLT	Prevents congestion using a color-based early drop mechanism. Schedules queues (PQ scheduling) based on priorities.	The color-based early drop mechanism ensures that all green packets of a priority are transmitted. PQ scheduling ensures that packets with a higher priority preferentially occupy bandwidth resources.

NOTE
One traffic stream can belong to only one group; one group can contain a maximum of eight service flows; all service flows of a group should belong to the same xPON port.

HQoS Implementation Principle on an xPON Board

HQoS on an xPON board support the following scenarios:

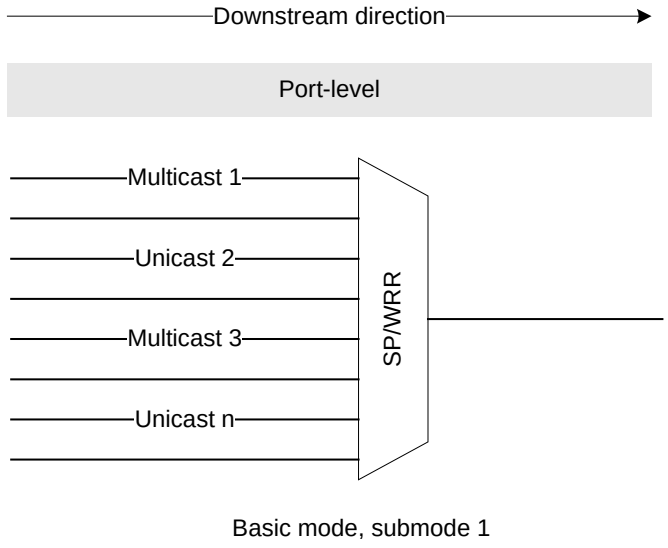
- Basic application
- Bandwidth priority
- Service priority

The following describes HQoS implementation principles in these 3 scenarios.

Basic Application Scenario

The basic mode inherits QoS functions on an xPON board. This mode has 2 sub-modes. Figure 14-53 shows the HQoS implementation principle in sub-mode 1 of the basic mode.

Figure 14-53 Sub-mode 1 of basic mode

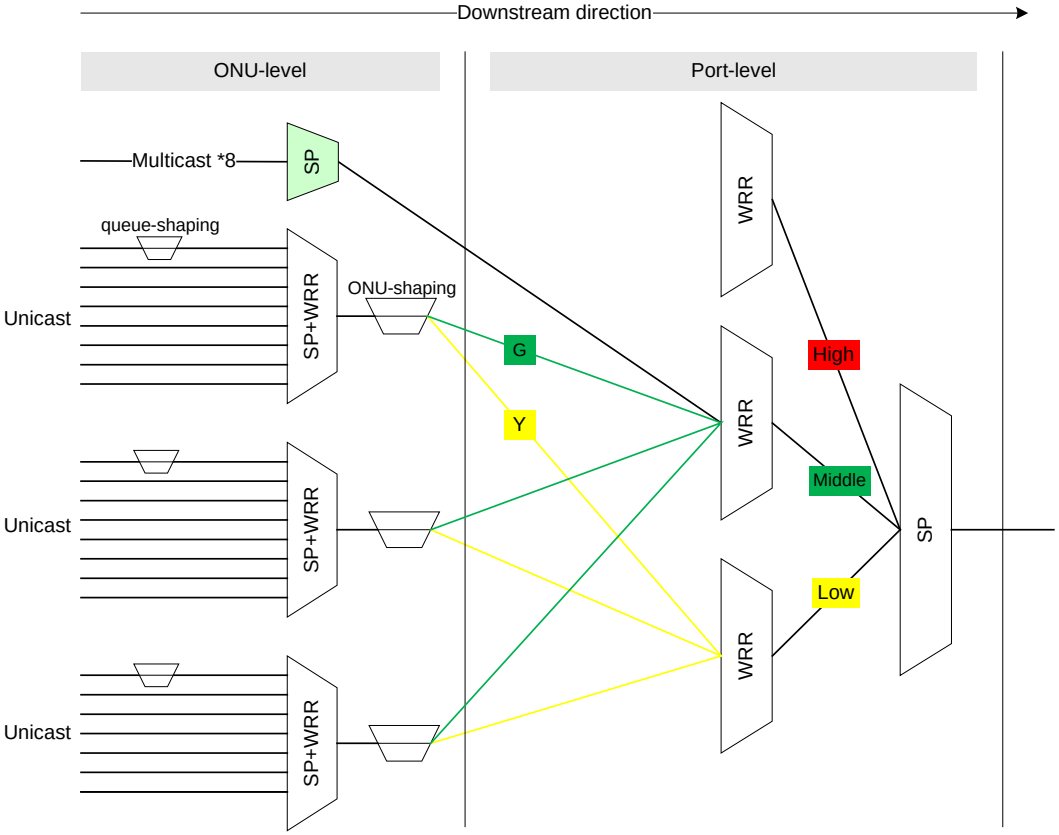


As shown in the figure, HQoS in this mode features the following:

- User unicast and multicast packets are enqueued on the PON port and are scheduled by SP and WRR by queue priorities.

When an xPON port has onu-shaping or queue-shaping, the HQoS mode is automatically switched to sub-mode 2 of the basic mode, as shown in Figure 14-54.

Figure 14-54 Sub-mode 2 of basic mode



Basic mode, submode 2

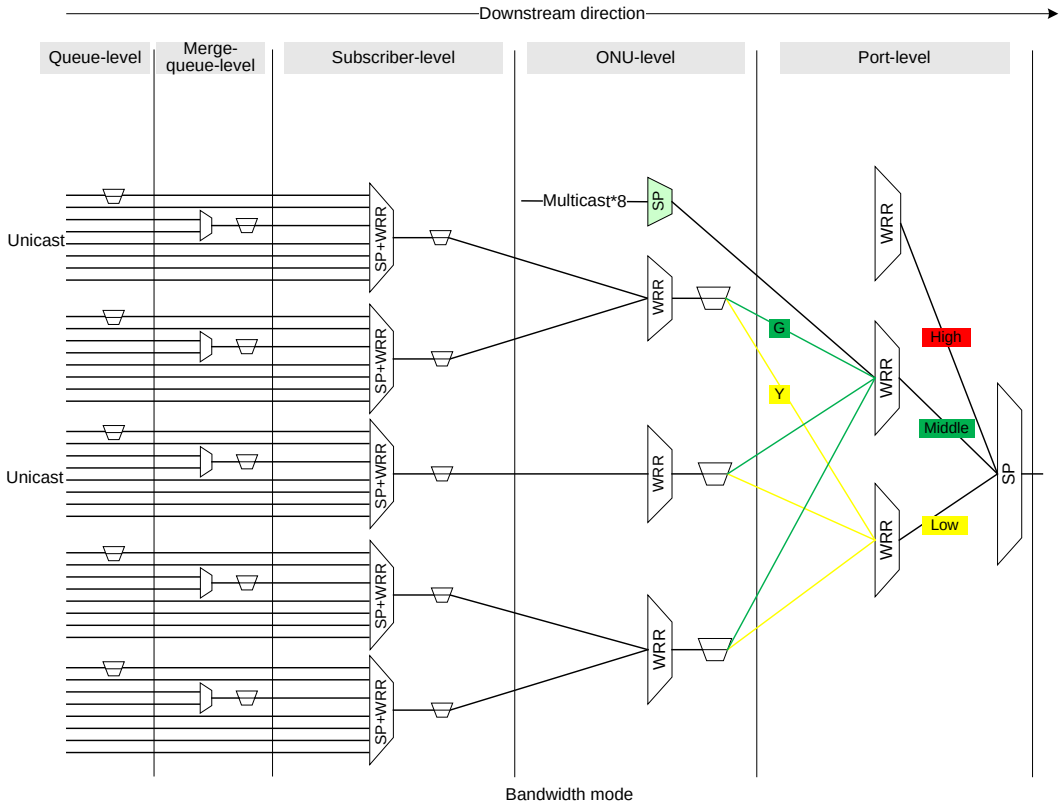
HQoS in this mode features the following:

- Multicast packets are enqueued on multicast queues.
- Unicast packets are enqueued on 8 priority queues of the ONU.
- Multicast duplication packets and unicast packets are scheduled by SP.
- 8 priority queues inside an ONU are scheduled by SP and WRR.
- Every ONU supports 8 queue shapers with only valid PIR.
- Every ONU supports a shaper with valid CIR and PIR.

Bandwidth Priority Scenario

In the bandwidth priority mode, the system preferentially assures bandwidths at all hierarchies for fair user usage, as shown in Figure 14-55.

Figure 14-55 Bandwidth priority mode



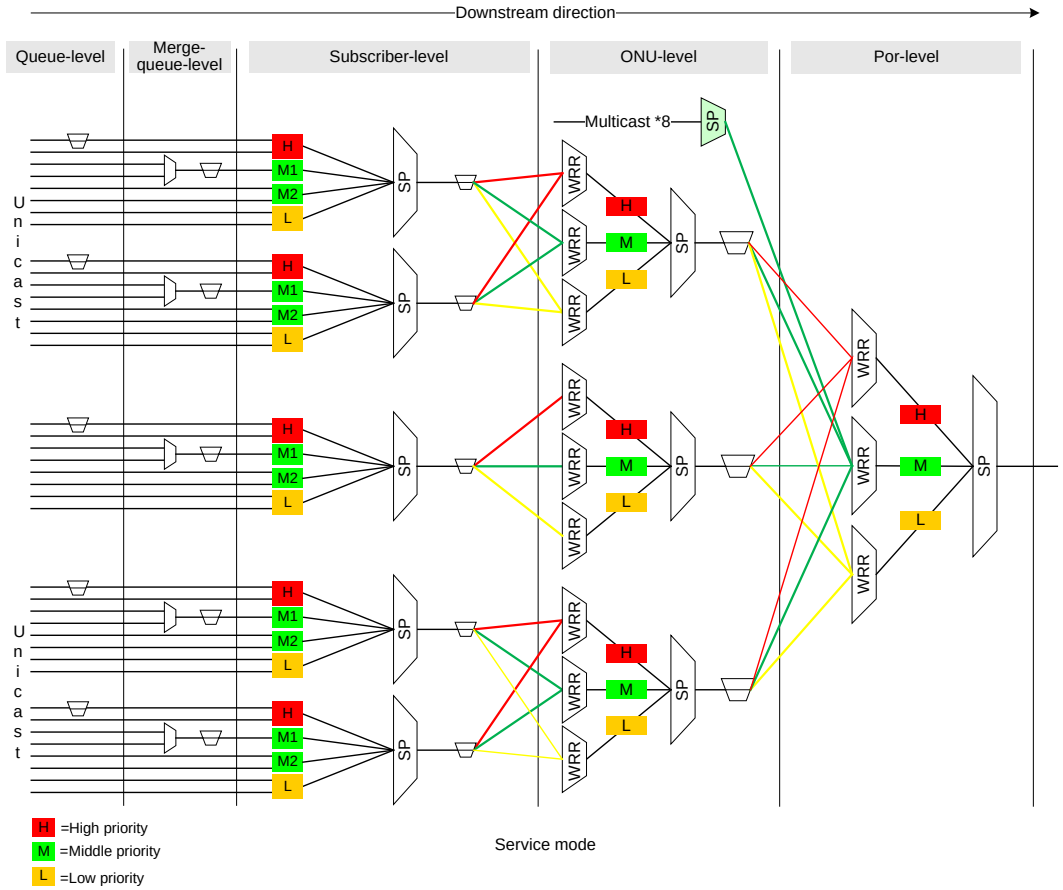
HQoS in this mode features the following:

- Multicast packets are enqueued on multicast queues.
- Unicast packets are enqueued on 8 priority queues.
- (Multicast duplication packets, green ONU packets, and) yellow ONU packets are scheduled by SP on a PON port.
- Every ONU supports a shaper with valid CIR and PIR. Green packets with CIR are scheduled between ONUs by RR and yellow packets with EIR (PIR-CIR) are scheduled between ONUs by WRR (weights are calculated by EIR).
- Users under an ONU are scheduled by WRR. Green packets with CIR are scheduled between users by RR and yellow packets with EIR (PIR-CIR) are scheduled between users by WRR (weights are calculated by EIR).
- Every user supports a shaper with valid CIR and PIR.
- Every user supports a merged queue shaper with only valid PIR.
- 8 priority queues inside users are scheduled by SP and WRR.
- 8 priority queues for users support queue shaper with only valid PIR.

Service Priority Scenario

In the service priority scenario, the system preferentially assures services with higher priorities for service scheduling, as shown in Figure 14-56.

Figure 14-56 Service priority mode



HQoS in this mode features the following:

- Multicast packets are enqueued on multicast queues.
- Unicast packets are enqueued on 8 priority queues.
- Packets with high priority, (multicast duplication packets and packets with medium priority, and) packets with low priority are scheduled by SP on a PON port.
- Every ONU supports a shaper with valid CIR and PIR.
- Packets having the same priority of different ONUs are scheduled by WRR.
- Packets having the same priority of different users under ONUs are scheduled by WRR (weights are calculated by PIR). Priorities high, medium, and low are supported for priority groups.
- Every user supports a shaper with only valid PIR.
- Every user supports a merged queue shaper with only valid PIR.
- 8 priority queues for users are scheduled by SP.
- 8 priority queues for users support queue shaper with only valid PIR.

14.12.8 Networking Application

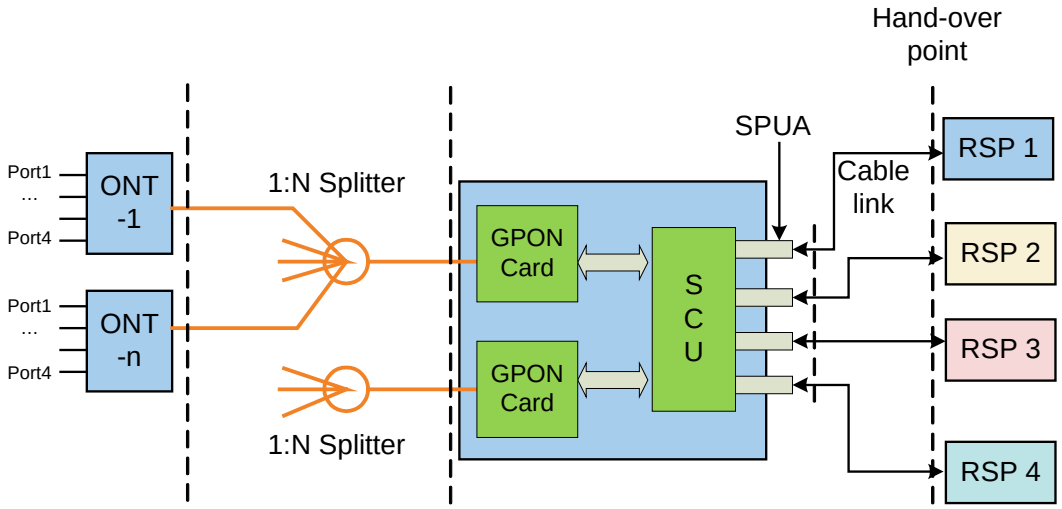
Networking Application of HQoS based on Port+VLAN

Figure 14-57 shows typical networking of hierarchical quality of service (HQoS) based on Port+VLAN. When Open Access is deployed, the carrier is only responsible for the network between the optical line terminal (OLT) and optical network terminal or ONT (including the ODN), while retail service providers (RSPs) are responsible for devices connected to the OLT and home area devices (an OTN provides 4 Ethernet UNI interfaces and each interface can be connected to one home device). The carrier network and RSP network are therefore interconnected at the uplink interface of the OLT (the uplink interface is provided by the SPUA board in this topic).

Generally, RSPs use different physical uplink interfaces. Sometimes, one RSP may use multiple physical uplink interfaces. If required, link convergence can be used to converge multiple uplink ports into one logical channel.

Since RSPs use different physical interfaces, each RSP can use all VLAN IDs, and duplicated VLAN IDs can be used for different RSPs.

Figure 14-57 Networking application of HQoS based on Port+VLAN



With the preceding networking, services that a carrier provides for RSPs cover the following types and the services address the following requirements:

- Voice services with symmetric uplink and downlink bandwidth, with CIR, but without PIR (voice services have the highest priority)
- Data services with symmetric uplink and downlink bandwidth, with CIR, but without PIR
- Data services with asymmetric uplink and downlink bandwidth (usually downlink bandwidth is far higher than the uplink bandwidth), with downlink PIR but without uplink PIR
- An ONT user port can carry traffic with multiple priorities:
  - Packets with a higher priority can preferentially occupy CIR resources. If there are remaining CIR resources, packets with a lower priority can be processed.

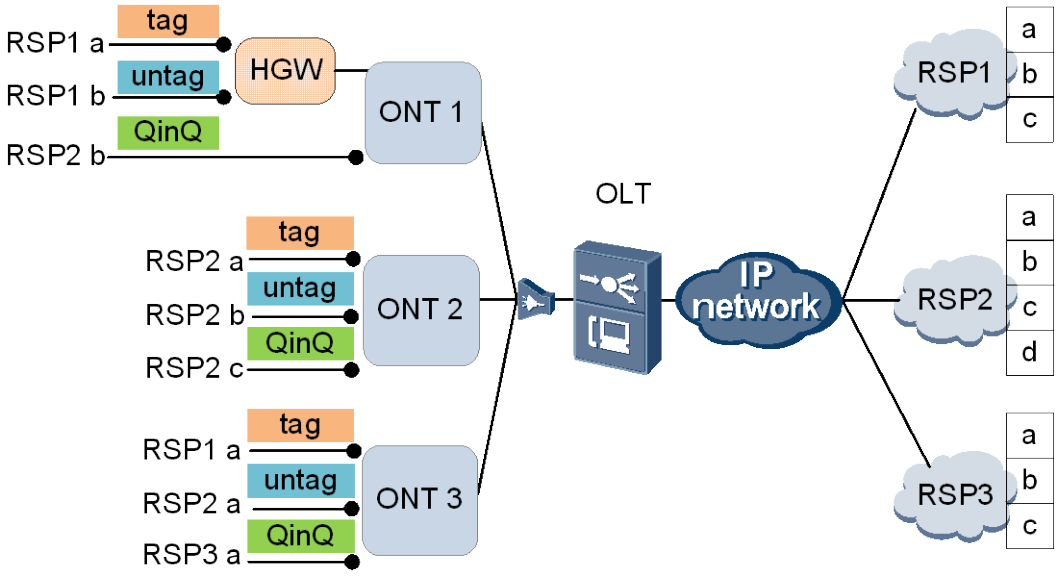
  - If the packet rate exceeds the CIR, packets with a higher priority can preferentially occupy PIR resources.
  - If there are remaining PIR resources, packets with a lower priority can be processed.
- For each type of data service, the CIR must be guaranteed. If required, PIR resources can be used in a certain proportion according to the network congestion status.

The preceding requirements can be well addressed by SPUA-based HQoS.

Networking Application of CAR-Group-based HQoS

As shown in Figure 14-58, users under the same ONT may belong to different RSPs. Different services of the same user are mapped to different service flows on the OLT. Given that different service CIRs/PIRs are guaranteed, the total bandwidth of each RSP needs to be ensured and each service should be allowed to occupy the total bandwidth when a burst occurs in the traffic. To put it simply, rate limitation needs to be performed on the RSP. To do so, a group based on all service flows of an RSP can be created, and then the total bandwidth of a user can be limited by limiting the bandwidth of the group. Such is a typical application of CAR group.

Figure 14-58 Networking application of HQoS based on a CAR group



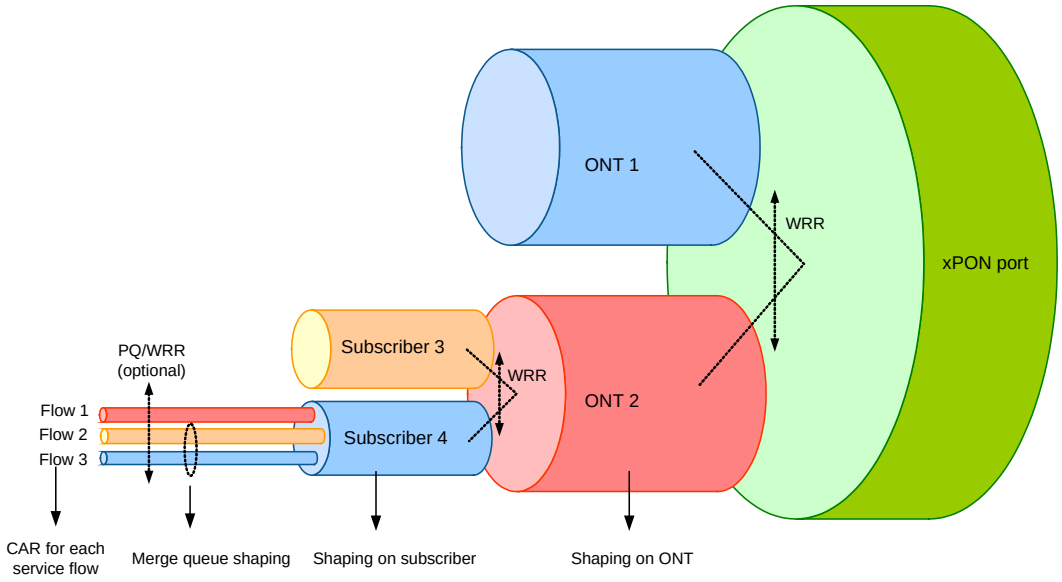
Networking Application of xPON-board-based HQoS

Figure 14-59 shows an HQoS application in the bandwidth priority mode of an xPON port. One xPON port connects to multiple ONTs, every ONT connects to multiple users, and every user is provisioned with multiple services. In this scenario, the bandwidth priority mode is used to ensure bandwidths of services at different hierarchies.

- CIR and EIR (PIR-CIR) permitted on different ONTs are allocated by preset weights.
- CIR and EIR (PIR-CIR) permitted by different users on an ONT are allocated by preset weights.
- WRR scheduling is used between users and between ONTs.

This application preferentially ensures user bandwidths for fair usage between users. It is applicable to the scenario in which the user priority is not a sensitive factor. In this case, when a network congestion occurs, the bandwidth plays an important role and users who subscribe to higher bandwidths have the high scheduling weight.

Figure 14-59 Networking application of xPON-board-based HQoS



14.12.9 Reference Standards and Protocols

The following lists the reference standards and protocols of this feature:

RFC2698: A Two Rate Three Color Marker

14.12.10 Configuring HQoS

The hierarchical QoS (HQoS) is a QoS technology that controls user traffic on a port with finer granularity and also schedules services of a user based on the service priority. This topic describes the configuration of HQoS.

Configuring HQoS Based on Port+VLAN

This topic describes how to configure the Port+VLAN-based HQoS to ensure different CIRs and PIRs for data with different priorities in the private line services (that is, data with a higher priority preferentially occupies the bandwidth).

Prerequisites

HQoS networking based on Port+VLAN is configured.

Context

Configuration method:

1. Configure a traffic profile.

2. Use the traffic profile to limit the rate of service flows for HQoS users.
3. Limit the rate of HQoS users based on the CAR threshold and mark colors of packets having different priorities.
4. Limit rates of HQoS user groups (port+SVLAN) based on colors.

Precaution:

HQoS is implemented through user configurations. Rate limitation on HQoS user groups is implemented using the single leaky bucket algorithm and bound TrTCM traffic profile. In the configured traffic profile, CIR = PIR; otherwise, PIR prevails for rate limitation. Ensure that PIR ≥ Total CIR bandwidths of all HQoS users.

Procedure

Configure a traffic profile.

Run the **traffic table ip** command to configure the CIR and PIR of a service, and priority copy policy and enqueuing policy of packets.

Step 1 Configure rate limitation on Layer 2 service flows (HQoS).

Run the **service-port** command to configure an HQoS user and use the traffic profile to limit the rate of this user.

Step 2 Limit the rate of the HQoS user based on the CAR threshold and mark packet colors.

Run the **car-threshold** command to configure drop thresholds for packets having different priorities and mark yellow and green for packets.

Step 3 Limit the rate of an HQoS user group based on port+VLAN.

Run the **car-port *portid* vlan** command to specify a traffic profile to limit the rate of an HQoS user group.

NOTE
An HQoS user group can be considered as a collection of users on a specified port within a specified VLAN range.

- **inbound** *ip-traffic-table-index*: Sets the traffic profile index for the packet from the outside of the device to the inside of the device.
- **outbound** *ip-traffic-table-index*: Sets the traffic profile index for the packet from the inside of the device to the outside of the device.

Step 4 Query the configuration results.

1. Run the **display traffic table ip** command to query the traffic profile configurations.
2. Run the **display service-port** command to query the HQoS user configurations.
3. Run the **display car-threshold** command to query drop thresholds for packets having different priorities.
4. Run the **display car-port** command to query rate limitation configurations of an HQoS user group.

----End

Example

In an open access network, a retail service provider (RSP) leases lines and bandwidths of a carrier and uses upstream port 0/2/0 on the SPUA board to provide the following data service package for FTTH users:

- Assured bandwidth 4 Mbit/s, symmetric upstream and downstream bandwidths, and permitted burst bandwidth 6 Mbit/s.
- User-side VLAN 20, network-side VLAN 100, trust user-side priority, and queue scheduling in egress queues by packet priorities.
- Priorities 0, 3, and 6 for the data service, and packets having the higher priority preferentially use the assured bandwidth and burst bandwidth.
- Packets having the lower priorities are dropped first when a congestion occurs on a port. Drop thresholds for packets having priorities 0, 3, and 6 are 15%, 50%, and 100%.
- Total upstream bandwidth of the data service on the port is 50 Mbit/s and no burst bandwidth is permitted; total downstream bandwidth is 200 Mbit/s and the permitted burst bandwidth is 250 Mbit/s.

Configure HQoS in this scenario:

```
//Configures the traffic profile for the data service of residential users by the RSP.
huawei(config)#traffic table ip index 8 cir 4096 pir 6144 priority user-cos
priority-policy tag-In-Package

//Configures the traffic profile for an RSP leased port by the carrier.
huawei(config)#traffic table ip index 9 cir 51200 priority user-cos priority-policy
tag-In-Package  //Indicates the downstream traffic profile.
huawei(config)#traffic table ip index 10 cir 204800 pir 256000 priority user-cos
priority-policy tag-In-Package  //Indicates the upstream traffic profile.

//Indicates an HQoS user (Layer 2 service flow).
huawei(config)#service-port 1 vlan 100 eth 0/2/0 multi-service user-vlan 20 rx-cttr
8 tx-cttr 8
huawei(config)#car-threshold cos0 15 cos3 50 cos6 100
huawei(config)#interface eth 0/2
huawei(config-if-eth-0/2)#car-port 0 vlan 100 inbound 10 outbound 9
```

After the configuration:

- Assured bandwidth 4 Mbit/s for the user when a network congestion occurs; burst bandwidth 6 Mbit/s for the user when the network is idle.
- Packets having different priorities share the bandwidth and packets are dropped from ones having the lower priorities when a network congestion occurs.
- Total bandwidth of the data service provided on the leased port by an RSP is at the most of the bandwidth allocated by the carrier.

Configuring HQoS Based on CAR Group

This topic describes how to configure HQoS based on CAR group for ensuring the bandwidth of each service of a user and the total bandwidth of the user.

Procedure

Configure CAR for user traffic streams.

In the global config mode, run the **service-port** command to specify a traffic profile for rate limitation of HQoS users.

Step 1 Configure CAR for an HQoS user.

1. Create a CAR group for a service port.

   In the global config mode, run the **car-group** command to create a CAR group for a service port and bind a traffic profile to the CAR group.

2. Add a service port to the CAR group.

   In the global config mode, run the **car-group add-member service-port** command to add a service port to the CAR group. Bandwidth of service ports in this CAR group is limited by the traffic profile bound to the CAR group.

Step 2 (Optional) Configure CAR for an HQoS user group.

In the ETH mode, run the **car-port***portid* **vlan** command to specify a traffic profile for rate limitation of an HQoS user group.

NOTE

An HQoS user group can be considered as a collection of users whose port IDs and VLAN IDs are within the port+VLAN range specified by this command.

- **inbound** *ip-traffic-table-index*: Sets the index of the traffic profile for packets transmitted from the outside of a device to the inside of the device.
- **outbound** *ip-traffic-table-index*: Sets the index of the traffic profile for packets transmitted from the inside of a device to the outside of the device.

----End

Example

Assume that the maximum bandwidth of a user is 5 Mbit/s. To configure 2 Mbit/s Internet access service, non-rate-limited voice service, and 4 Mbit/s multicast service for the user, do as follows:

```
huawei(config)#traffic table ip index 8 cir 2048 priority 1 priority-policy
tag-In-Package
huawei(config)#traffic table ip index 9 cir off priority 6 priority-policy
tag-In-Package
huawei(config)#traffic table ip index 10 cir 4096 priority 4 priority-policy
tag-In-Package
huawei(config)#traffic table ip index 20 cir 5120 priority 6 priority-policy
tag-In-Package
huawei(config)#service-port 1 vlan 100 gpon 0/1/1 ont 1 gemport 1 multi-service user-vlan
10 rx-cttr 8 tx-cttr 8
huawei(config)#service-port 2 vlan 200 gpon 0/1/1 ont 1 gemport 2 multi-service user-vlan
20 rx-cttr 9 tx-cttr 9
huawei(config)#service-port 3 vlan 300 gpon 0/1/1 ont 1 gemport 3 multi-service user-vlan
30 rx-cttr 10 tx-cttr 10
huawei(config)#car-group 1 inbound traffic-table index 20 outbound traffic-table index
20
huawei(config)#car-group 1 add-member service-port 1-3
```

Configuring HQoS for an xPON Board

In an xPON access scenario, select a proper HQoS mode according to actual QoS requirements.

Prerequisites

Supported xPON boards: H807GPBH, H805GPFD, and H801XGBD.

Context

xPON boards support HQoS applications in the following exclusive scenarios:

- Basic application scenario: Inherits the original QoS function. It is the default mode.
- Bandwidth priority scenario: The system preferentially assures bandwidths at all hierarchies for fair user usage.
- Service priority scenario: The system preferentially assures services with higher priorities for service scheduling.

For detailed application scenarios, see HQoS Implementation Principle on an xPON Board.

Procedure

- **Configure HQoS in a basic application scenario.**
  a. In xPON mode. run the **hqos mode** ***basic*** command to set the HQoS mode of an xPON board to basic.
  b. Run the **display hqos mode** command to query the HQoS mode of a board.
- **Configure HQoS in a bandwidth priority scenario.**
  a. Set the HQoS mode to **bandwidth**.

     In xPON mode. run the **hqos mode** ***bandwidth*** command to set the HQoS mode of an xPON board to bandwidth priority.
  b. Configure rate limitation on an ONT.

     Run the **traffic-limit ont** command to limit the traffic of downstream packets on a specified ONT.
  c. Configure the HQoS user profile and its attributes.
     i. Run the **subscriber profile** command to create an HQoS user profile.
     ii. Run the **subscriber shaping** command to limit the rate of an HQoS user.
  d. (Optional) Configure the merged queue profile of HQoS users and its attributes.
     i. Run the **merge queue profile** command to create a merged queue profile and bind queues.
     ii. Run the **merge queue bind** command to bind a merged queue profile.
     iii. Run the **merge queue shaping** command to set rate limitation on the merged queue.
  e. Add an HQoS user and limit the rate of this user.
     i. Run the **subscriber add** command to bind a user to the HQoS user profile and limit the rate of this user.
     ii. Run the **subscriber member add** command to add a service port as an HQoS user.
  f. Query the configuration.

- Run the **display hqos mode** command to query the HQoS mode of a board.
- Run the **display subscriber** command to query the HQoS user information.
- (Optional) Run the **display merge queue profile** command to query the merged queue information about an HQoS user.

- **Configure HQoS in a service priority scenario.**

NOTE

HQoS configuration in this scenario is the same as that in the bandwidth priority scenario. The only difference is that HQoS mode in this scenario is set to **service**.

a. Set the HQoS mode to **service**.

   In xPON mode. run the **hqos mode** ***service*** command to set the HQoS mode of an xPON board to service priority.

b. Configure rate limitation on an ONT.

   Run the **traffic-limit ont** command to limit the traffic of downstream packets on a specified ONT.

c. Configure the HQoS user profile and its attributes.

   i. Run the **subscriber profile** command to create an HQoS user profile.

   ii. Run the **subscriber shaping** command to limit the rate of an HQoS user.

d. (Optional) Configure the merged queue profile of HQoS users and its attributes.

   i. Run the **merge queue profile** command to create a merged queue profile and bind queues.

   ii. Run the **merge queue bind** command to bind a merged queue profile.

   iii. Run the **merge queue shaping** command to set rate limitation on the merged queue.

e. Add an HQoS user and limit the rate of this user.

   i. Run the **subscriber add** command to bind a user to the HQoS user profile and limit the rate of this user.

   ii. Run the **subscriber member add** command to add a service port as an HQoS user.

f. Query the configuration.

   - Run the **display hqos mode** command to query the HQoS mode of a board.
   - Run the **display subscriber** command to query the HQoS user information.
   - (Optional) Run the **display merge queue profile** command to query the merged queue information about an HQoS user.

----End

Example

The following is an example of the configurations used to plan 3 ONTs under GPON port 0/1/0:

- Set the bandwidth priority mode.
- Configure a user for ONT 1 (ONT ID 1).
  - Configure 2 service flows: the rate of one service flow with index 100 and priority 4 is limited to 600 Mbit/s and that of the other one with index 101 and priority 0 is limited to 150 Mbit/s.

  - Limit the shaping rate of users to 800 Mbit/s (traffic profile index 10, CIR 400 Mbit/s, and PIR 800 Mbit/s).
  - Limit the shaping rate of ONT 1 to 1 Gbit/s (traffic profile index 11, CIR 400 Mbit/s, and PIR 1 Gbit/s).
- Configure ONT 2 (ONT ID 2) in the same way as ONT 1 (the indexes of the two service flows are 102 and 103).
- Configure 2 users for ONT 3 (ONT ID 3) under 2 GE ports on the ONT.
  - Configure 2 service flows for user 1 (ID 2): the rate of one service flow with index 102 and priority 4 is limited to 600 Mbit/s and that of the other one with index 103 and priority 0 is limited to 150 Mbit/s.
  - Limit the shaping rate of user 1 to 800 Mbit/s (traffic profile index 12, CIR 600 Mbit/s, and PIR 800 Mbit/s).
  - Configure 2 service flows for user 2 (ID 3): the rate of one service flow with index 104 and priority 4 is limited to 600 Mbit/s and that of the other one with index 105 and priority 0 is limited to 150 Mbit/s.
  - Limit the shaping rate of user 2 to 1 Gbit/s (CIR 400 Mbit/s and PIR 1 Gbit/s).
  - Limit the shaping rate of ONT 3 to 2 Gbit/s (traffic profile index 13, CIR 1.2 Gbit/s, and PIR 2 Gbit/s).

The configuration procedures are as follows:

```
  //Configure the HQoS mode.
huawei(config)#interface gpon 0/1
huawei(config-if-gpon-0/1)#hqos mode bandwidth
huawei(config-if-gpon-0/1)#quit

..//Cofnigure the traffic profile.
huawei(config)#traffic table ip index 10 cir 409600 pir 819600 priority user-cos 0
priority-policy tag-In-package
huawei(config)#traffic table ip index 11 cir 409600 pir 1024000 priority user-cos 0
priority-policy tag-In-package
huawei(config)#traffic table ip index 12 cir 614400 pir 819600 priority user-cos 0
priority-policy tag-In-package
huawei(config)#traffic table ip index 13 cir 1228800 pir 2048000 priority user-cos 0
priority-policy tag-In-package

  //Configure the profile for HQoS users under ONT 1 and ONT 2.
huawei(config)#subscriber profile profile-id 1
huawei(config-subscriber-profile-1)#subscriber shaping outbound traffic-table index
10
huawei(config-subscriber-profile-1)#quit

  //Add HQoS users under ONT 1 and ONT 2.
huawei(config)#subscriber add 0/1/0 ont 1 1 profile-id 1
huawei(config)#subscriber member add 0/1/0 ont 1 1 service-port 100
huawei(config)#subscriber member add 0/1/0 ont 1 1 service-port 101
huawei(config)#subscriber add 0/1/0 ont 2 1 profile-id 1
huawei(config)#subscriber member add 0/1/0 ont 2 1 service-port 102
huawei(config)#subscriber member add 0/1/0 ont 2 1 service-port 103

  //Limit rates of ONT 1 and ONT 2.
huawei(config)#traffic-limit ont 0/1/0 1 down-stream traffic-table index 11
```

```
huawei(config)#traffic-limit ont 0/1/0 2 down-stream traffic-table index 11

 //Configure the profile for HQoS user 1 under ONT 3.
huawei(config)#subscriber profile profile-id 2
huawei(config-subscriber-profile-2)#subscriber shaping outbound traffic-table index
12
huawei(config-subscriber-profile-2)#quit

  //Add HQoS user 1 under ONT 3.
huawei(config)#subscriber add 0/1/0 ont 3 2 profile-id 2
huawei(config)#subscriber member add 0/1/0 ont 2 2 service-port 102
huawei(config)#subscriber member add 0/1/0 ont 2 2 service-port 103

  //Configure the profile for HQoS 2 under ONT 3.
huawei(config)#subscriber profile profile-id 3
huawei(config-subscriber-profile-3)#subscriber shaping outbound traffic-table index
11
huawei(config-subscriber-profile-3)#quit

  //Add HQoS user 3 under ONT 3.
huawei(config)#subscriber add 0/1/0 ont 3 3 profile-id 3
huawei(config)#subscriber member add 0/1/0 ont 3 3 service-port 104
huawei(config)#subscriber member add 0/1/0 ont 3 3 service-port 105

  //Limit the rate of ONT 3.
huawei(config)#traffic-limit ont 0/1/0 3 down-stream traffic-table index 13
```

Follow-up Procedure

Expected results:

- The CIR ratio of ONT 1, ONT 2, and ONT 3 is 1:1:3, and the total CIR of these 3 ONTs is 2.0 Gbit/s, which does not exceed the downstream bandwidth of a PON port.
- The EIR (PIR-CIR) ratio of ONT 1, ONT 2, and ONT 3 is 3:3:4, and the remaining bandwidth of a PON port can be used by EIR.
- ONT 1 permits bandwidth about 550 Mbit/s, containing CIR 400 Mbit/s for packets with priority 4 and EIR 150 Mbit/s for packets with priority 4.
- ONT 2 permits bandwidth about 550 Mbit/s, containing CIR 400 Mbit/s for packets with priority 4 and EIR 150 Mbit/s for packets with priority 4.
- ONT 3 permits bandwidth about 1400 Mbit/s:
  - The CIR ratio of user 1 and user 2 is 6:4 and EIR ratio is 1:3.
  - Packet bandwidth of user 1 is 700 Mbit/s, containing CIR 600 Mbit/s for packets with priority 4 and EIR 100 Mbit/s for packets with priority 0.
  - Packet bandwidth of user 2 is 700 Mbit/s, containing CIR 400 Mbit/s for packets with priority 4 and EIR 300 Mbit/s for packets with priority 4 (200 Mbit/s) and with priority 0 (100 Mbit/s).

14.13 End-to-End QoS

14.13.1 FTTH End-to-End QoS Policy

The FTTH end-to-end (E2E) QoS solution uses the differentiated service (DiffServ) model. In this solution, QoS parameters are carried in each packet header to ensure QoS guarantees. The following describes upstream and downstream FTTH E2E QoS applications.

Upstream E2E QoS

Upstream E2E QoS is shown in the Figure 14-60 and Table 14-17.

1. QoS policy on the ONT.
   a. An ONT classifies traffic based on user ports or user-side VLAN IDs to distinguish services.
   b. Re-marks the 802.1p priority for service packets.
   c. Based on packet priorities, the ONT arranges packets into different priority queues by using PQ scheduling. This ensures that services with higher QoS requirements are forwarded preferentially.
   d. Limits the upstream traffic on user-side ports. (optional).
2. QoS policy on the OLT.
   a. Classifies traffic based on VLAN IDs and the 802.1p priority
   b. Trusts the user-side priority or re-marks the 802.1p priority for service packets.
   c. Implements priority scheduling and congestion management based on the 802.1p priority.
   d. Implements DBA to limit the ONUs upstream bandwidth.
   e. limits upstream traffic bandwidth based on service flows (optional).
3. The layer 2 metropolitan area network (MAN) between the OLT and the BRAS/service router (SR) implements priority scheduling based on the 802.1p priority of service packets.
4. Traffic classification is enabled on the ingress of the BRAS/SR to identify 802.1p values of different packets and re-mark packets according to the upper network.
   - When the upper-layer network is a native IP Layer 3 network, the BRAS/SR maps the 802.1p priority to the ToS/DSCP value in the IP header.
   - When the upper-layer network is an MPLS VPN, the BRAS/SR maps the 802.1p priority to MPLS EXP.

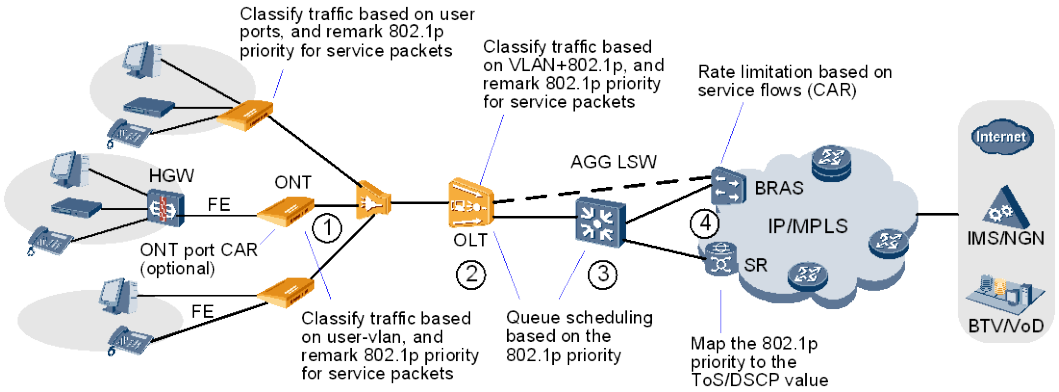
   The BRAS/SR can also limit bandwidth for upstream user traffic based on service flows (optional).
5. Upper-level networks of the BRAS/SR implement traffic classification, congestion management, and congestion avoidance based on ToS/DSCP in the IP header or the MPLS EXP priority.

Table 14-17 Upstream E2E QoS policy

Device	QoS policy			
	Traffic classification	Priority processing	Traffic policing	Queue scheduling

Device	QoS policy			
	Traffic classification	**Priority processing**	**Traffic policing**	**Queue scheduling**
ONT	• Based on user ports • Based on user-side VLAN IDs	Re-marks the 802.1p priority	Limits the upstream traffic on user-side ports. (optional).	PQ schedule based on 802.1p priority
OLT	VLAN+802.1p priority	• Trusts the user-side priority • Re-marks the 802.1p priority	Implements DBA to limit the ONUs upstream bandwidth. limits upstream traffic bandwidth based on service flows (optional).	PQ schedule based on 802.1p priority
L2 MAN	N/A	N/A	N/A	PQ schedule based on 802.1p priority of service packets
BRAS/SR	Based on VLAN	• Maps the 802.1p priority to the ToS/DSCP value in the IP header when upper-layer is IP L3 network. • Maps the 802.1p priority to MPLS EXP when upper-layer is MPLS VPN.	limits bandwidth for upstream user traffic based on service flows (optional)	N/A
L3 network (IP/MPLS)	IP network: based on ToS/DSCP MPLS network: based on MPLS EXP	Trusts priority of packets	N/A	PQ

Figure 14-60 FTTH E2E QoS networking



Downstream E2E QoS

Downstream E2E QoS is shown in the Figure 14-60 and Table 14-18.

1. The packets of the IPTV, VoIP, and broadband Internet services are assigned different DSCP/MPLS EXP values based on priority requirements using related egress switches or provider edge (PE) devices. Core networks and backbone networks perform priority-based queue scheduling according to DSCP/MPLS EXP.
2. Downstream packets are transmitted to edge BRAS/SR of the Layer 3 network through core networks and backbone networks. The BRAS/SR maps IP DSCP or MPLS EXP of packets to the VLAN 802.1p priority. The BRAS/SR can also limit bandwidth for downstream user traffic based on service flows (optional).
   - For broadband Internet services in PPPoE mode, the BRAS controls the access bandwidth of each broadband user based on the rate authorized by the RADIUS server. For example, the rate can be 512 kbit/s, 1 Mbit/s, or 2 Mbit/s.
   - Due to traffic jitter of the server or network, queue shaping is necessary for IPTV services on service routers.
3. Metropolitan networks and Layer 2 access networks downstream to the BRAS/SR implement priority scheduling based on the packet 802.1p priority.
4. The OLT classifies traffic based on VLAN IDs and the 802.1p priority, and it trusts the network-side priority or re-marks the 802.1p priority for service packets. It implements priority scheduling and congestion management based on the 802.1p priority. It can limit downstream traffic bandwidth based on service flows (optional).
5. The ONT trusts the network-layer 802.1p priority and implements priority scheduling based on the 802.1p priority. The ONT can also limit downstream traffic bandwidth based on user-side ports (optional).

Table 14-18 Downstream E2E QoS policy

Device	QoS policy			
	Traffic classification	Priority processing	Traffic policing	Queue scheduling
L3 netwo	N/A	Assign different priority values	N/A	PQ

Device	QoS policy			
	Traffic classification	**Priority processing**	**Traffic policing**	**Queue scheduling**
rk (IP/MPLS)		based on service types and networks. • IP network: assign different IP ToS/DSCP value • MPLS network: assign different MPLS EXP value		
BRAS/SR	N/A	• Maps the ToS/DSCP value to 802.1p priority when upper-layer is IP L3 network. • Maps the MPLS EXP to 802.1p priority when upper-layer is MPLS network.	Limits bandwidth for downstream user traffic based on service flows (optional). • HSI service in PPPoE mode: CAR • IPTV service: queue shaping	N/A
L2 MAN	N/A	N/A	N/A	PQ schedule based on 802.1p priority of service packets
OLT	VLAN+802.1p priority	• Trusts the network-layer 802.1p priority • Re-marks 802.1p priority	Limits downstream traffic bandwidth based on service flows (optional)	PQ schedule based on 802.1p priority
ONT	N/A	Trusts the network-layer 802.1p priority	Limits downstream traffic bandwidth based on user-side ports (optional)	PQ schedule based on 802.1p priority

14.13.2 FTTB/FTTC End-to-End QoS Policy

The FTTB/FTTC E2E QoS solution uses the DiffServ model. In this solution, QoS parameters are carried in each packet header to ensure QoS guarantees. The following describes upstream and downstream E2E QoS applications for FTTB/FTTC.

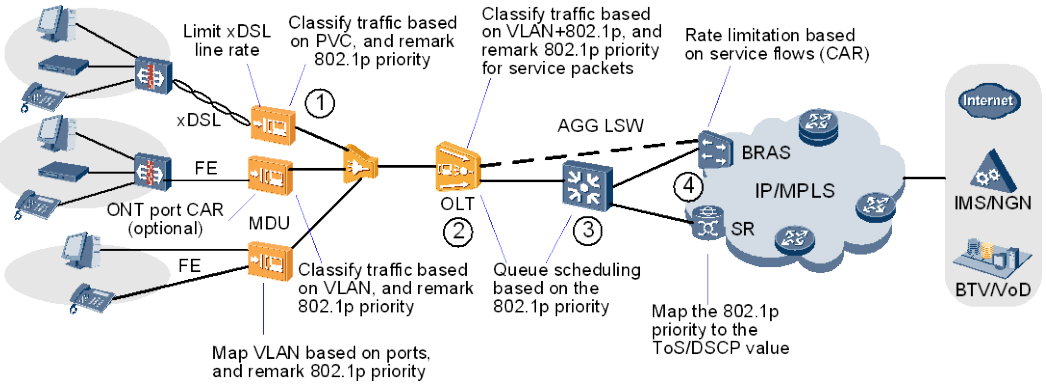
Upstream E2E QoS

Figure 14-61 shows the upstream E2E QoS.

1. QoS policy on the MDU.
   a. An MDU classifies traffic based on the user port or user-side VLAN IDs/PVC (xDSL access) to distinguish services
   b. Re-marks the 802.1p priority for service packets.
   c. Based on packet priorities, the MDU arranges packets into different priority queues by using PQ scheduling. This operation ensures that services with high QoS are forwarded preferentially.
   d. Limits the upstream traffic on user-side ports.(optional).
2. QoS policy on the OLT.
   a. Classifies traffic based on VLAN IDs and the 802.1p priority
   b. Trusts the user-side priority or re-marks the 802.1p priority for service packets.
   c. Implements priority scheduling and congestion management based on the 802.1p priority.
   d. Implements DBA to limit the ONUs upstream bandwidth.
   e. limits upstream traffic bandwidth based on service flows (optional).
3. The layer 2 metropolitan area network (MAN) between the OLT and the BRAS/service router (SR) implements priority scheduling based on the 802.1p priority of service packets.
4. Traffic classification is enabled on the ingress of the BRAS/SR to identify 802.1p values of different packets and re-mark packets according to the upper network.
   - When the upper-layer network is a native IP Layer 3 network, the BRAS/SR maps the 802.1p priority to the ToS/DSCP value in the IP header.
   - When the upper-layer network is an MPLS VPN, the BRAS/SR maps the 802.1p priority to MPLS EXP.

   The BRAS/SR can also limit bandwidth for upstream user traffic based on service flows (optional).
5. Upper-level networks of the BRAS/SR implement traffic classification, congestion management, and congestion avoidance based on ToS/DSCP in the IP header or the MPLS EXP priority.

Figure 14-61 FTTB/FTTC E2E QoS networking



Downstream E2E QoS

Figure 14-61 shows the downstream E2E QoS.

1. The packets of the IPTV, VoIP, and broadband Internet services are assigned different DSCP/MPLS EXP values based on priority requirements using related egress switches or provider edge (PE) devices. Core networks and backbone networks perform priority-based queue scheduling according to DSCP/MPLS EXP.
2. Downstream packets are transmitted to edge BRAS/SR of the Layer 3 network through core networks and backbone networks. The BRAS/SR maps IP DSCP or MPLS EXP of packets to the VLAN 802.1p priority. The BRAS/SR can also limit bandwidth for downstream user traffic based on service flows (optional).
   - For broadband Internet services in PPPoE mode, the BRAS controls the access bandwidth of each broadband user based on the rate authorized by the RADIUS server. For example, the rate can be 512 kbit/s, 1 Mbit/s, or 2 Mbit/s.
   - Due to traffic jitter of the server or network, queue shaping is necessary for IPTV services on service routers.
3. Metropolitan networks and Layer 2 access networks downstream to the BRAS/SR implement priority scheduling based on the packet 802.1p priority.
4. The OLT classifies traffic based on VLAN IDs and the 802.1p priority, and it trusts the network-side priority or re-marks the 802.1p priority for service packets. It implements priority scheduling and congestion management based on the 802.1p priority. It can limit downstream traffic bandwidth based on service flows (optional).
5. The MDU trusts the network-layer 802.1p priority and implements priority scheduling based on the 802.1p priority. The user-side port on the MDU limits bandwidth for downstream traffic (optional).

14.13.3 QoS Solution for FTTH

Basic Principles of QoS for FTTH

1. The ingress (upstream port on ONT/downstream port on OLT) of an FTTx access network classifies traffic and re-marks the 802.1p priority for packets. The access device (MA5600T/MA5603T/MA5608T) implements congestion control in PQ scheduling mode based on the 802.1p priority.
2. Service packets on an ONT are differentiated based on VLAN IDs.

- In a GPON system, GEM port mapping is implemented based on VLAN IDs to specify whether different traffic packets are transmitted to different GEM ports or to the same GEM port. Each ONT uses a T-CONT.

3. DBA profiles are configured based on users' bandwidth types. During service provisioning, relevant DBA profiles are selected based on the bandwidth types that users apply for. The recommended DBA type is type 3 (assured bandwidth+maximum bandwidth). Upstream priority of the T-CONT for an ONT follows 802.1p-based priority scheduling.
4. It is recommended to use the same VLAN configuration for all ONTs and implement VLAN translation on the OLT. This operation ensures that the same type of ONTs having the same bandwidth types use the same line profile and service profile.
5. It is recommended to limit service flow bandwidth on the BRAS or service router (SR) but not on the access device.

NOTE

The ONT can be configured to limit bandwidth for user-side ports, and the OLT can be configured to rate limit service flows based on traffic profiles.

Recommended QoS Plan for FTTH

The following table provides details about the recommended QoS plan for FTTH with respect to traffic classification, priority marking, and queue scheduling policies.

Service Type	802.1p Priority	Queue Scheduling Method	OLT Queue ID (8 Queues)	ONT Queue ID	
				When Eight Queues Are Supported	When Four Queues Are Supported
Management service	6	PQ	6	6	3
VoIP service	5	PQ	5	5	2
IPTV service	4	PQ	4	4	2
Internet access service	0	PQ	0	0	0

Note: The service priorities in this table are recommended values only. Configure the service priorities according to the operator's actual network plan.

The following table provides details about the recommended QoS plan for FTTH with respect to traffic policing and DBA policies.

Items	Management service	Internet access service	VoIP service	IPTV service
GEM port	11 [Remark 1]	12	13	14

Items	Management service	Internet access service	VoIP service	IPTV service
TCONT	All services share a T-CONT.			
DBA type	Type 3 (DBA profile: assured bandwidth + maximum bandwidth. Users are allowed to preempt the bandwidth on condition that the users' assured bandwidth is guaranteed. However, the total bandwidth cannot exceed the maximum bandwidth.)			
DBA bandwidth planning	Configure the DBA bandwidth according to the user's bandwidth package. The assured bandwidth is the maximum bandwidth required by management packets, VoIP, and IPTV upstream packets. The maximum bandwidth is larger than or equal to the maximum bandwidth that users apply.			
Rate limit on OLT downstream	No rate limit	Configure rate limit by a traffic profile as required. [Remark 2]	No rate limit	No rate limit
Rate limit on ONU upstream port	No rate limit			
Rate limit on ONU downstream port	No rate limit			

Remark 1: GEM port values depend on the planning of the service provider. Generally, however, use different GEM ports for different services.

Remark 2: Table 14-19 shows the reference service bandwidth of each service for each user.

Table 14-19 Reference service bandwidth of each service for each user

Service Type	Upstream Bandwidth	Downstream Bandwidth	Bandwidth Description
Internet access service	Determined based on the service package.	Determined based on the service package.	Available bandwidth of Internet access service = Committed bandwidth of the service package - VoIP bandwidth - IPTV bandwidth
VoIP service	200 kbit/s	200 kbit/s	The upstream bandwidth and the downstream bandwidth of VoIP service are symmetrical. The actual bandwidth is related to the coding format. This

Service Type	Upstream Bandwidth	Downstream Bandwidth	Bandwidth Description
			bandwidth is calculated for two POTS ports.
IPTV service (standard definition program)	/	2.5 Mbit/s per channel	IPTV service mainly occupies the downstream bandwidth. The actual bandwidth depends on the coding format, the picture in picture (PiP) information, 10% of the assured bandwidth for burst traffic, and the number of programs that can be concurrently watched by one user (in the case of multiple STBs). The upstream bandwidth is mainly used for transmitting IGMP packets, which requires only a little bandwidth. Therefore, the bandwidth occupied by IGMP packets can be ignored.
IPTV service (high definition program)	/	9.7 Mbit/s per channel	

Note:

- If the BRAS does not support rate limitation, OLTs can limit the rate for service flows by using traffic profiles.
- Different service packets on the ONT are distinguished by different VLAN IDs. The service packets are mapped to GEM ports based on VLAN IDs so that different service packets are transmitted to different GEM ports. Each GEM port (each service) can correspond to a T-CONT or all GEM ports share a T-CONT.
- The sum of the assured bandwidth of all ONTs connected to an OLT PON port and the fixed bandwidth of OMCI management channel should be smaller than the GPON upstream bandwidth. Some bandwidth must be reserved for future service expansion.

14.13.4 QoS Solution for FTTB/FTTC

Basic Principles of QoS for FTTB/FTTC

1. Plan different CoS priorities for different services. The advised priority order is as follows: management packets > VoIP > IPTV > HSI.
2. Classify traffic and re-mark 802.1p priority on the ingress (upstream port on MDU/downstream port on OLT) on the FTTx network. Implement congestion control based on 802.1p priority in PQ scheduling mode on the internal access network (MDU/OLT). For packets transmitted downstream from the network side of the OLT, the priority can be trusted. The priority does not need to be re-marked.
3. Service packets on MDUs are distinguished by 802.1p priority.
   - GEM port mapping is implemented in a GPON system based on 802.1p, so that different service packets enter different GEM ports. Each GEM port (each service) can correspond to a T-CONT or all GEM ports share a T-CONT.

NOTE
T-CONT 0 is still used exclusively for OMCI management.

4. DBA planning:
   - GPON MDUs sharing one T-CONT: Set different DBA profiles for different MDU types, and then choose corresponding DBA profiles in the service pre-provisioning phase based on the MDU type. Type 3 (assured bandwidth+maximum bandwidth) is recommended for a DBA profile. The T-CONT upstream priority for MDUs is based on 802.1p priority.
   - GPON MDUs using multiple T-CONTs: Allocate different DBA profiles for T-CONTs of each service.
5. Bandwidth limit: It is recommended to implement bandwidth limit of user service flows on the BRAS/SR, not on the access network. MDUs can be configured to limit the bandwidth on the user-side ports; MDUs/OLTs can be configured to limit the rate of service flows based on traffic profiles.
6. For xDSL access, an MxU supports bandwidth limit on physical links using the xDSL link traffic profile. Bandwidth limit on xDSL link is based on the fact that the channels on user-side ports do not differentiate multiple service PVCs. It is recommended that you differentiate services on the modem for users with large upstream traffic so that QoS can be ensured. The MA5821 and MA5822 do not support xDSL access.

Recommended QoS Plan for FTTB/FTTC

The following table provides details about the recommended QoS plan for FTTB/FTTC with respect to traffic classification, priority marking, and queue scheduling policies.

Service Type	802.1p Priority	Queue Scheduling Method	OLT Queue ID (8 Queues)	ONT Queue ID	
				When Eight Queues Are Supported	When Four Queues Are Supported
Management service	6	PQ	6	6	3
VoIP service	5	PQ	5	5	2
IPTV service	4	PQ	4	4	2
Internet access service	0	PQ	0	0	0

Note

- Different service packets are distinguished by different VLAN IDs. GEM ports are mapped based on 802.1p priorities for the GPON system.
- Service priorities in this table are for reference only. Configure the service priorities according to the operator's actual network plan.

The following table provides details about the recommended QoS plan for FTTB/FTTC with respect to traffic policing and DBA policies.

Items	Management service	Internet access service	VoIP service	IPTV service
GEM port	11 [Remark 1]	12	13	14
T-CONT	All services share a T-CONT.			
DBA type	Type 3 (DBA profile: assured bandwidth + maximum bandwidth. Users are allowed to preempt the bandwidth on condition that the users' assured bandwidth is guaranteed. However, the total bandwidth cannot exceed the maximum bandwidth.)			
DBA bandwidth planning	Configure the DBA bandwidth according to the user's bandwidth package. The assured bandwidth is the maximum bandwidth required by management packets, VoIP, and IPTV upstream packets. The maximum bandwidth is larger than or equal to the maximum bandwidth that users apply.			
Rate limit on OLT downstream	No rate limit	Configure rate limit by a traffic profile as required. [Remark 2]	No rate limit	No rate limit
Rate limit on ONU upstream port	Set ONU port rate limit or xDSL line rate limit as required. [Remark 2]			
Rate limit on ONU downstream port	Set ONU port rate limit or xDSL line rate limit as required. [Remark 2]			

Remark 1: GEM port values depend on the planning of the service provider. Generally, however, use different GEM ports for different services. The MA5821 and MA5822 do not support xDSL access.

Remark 2: Table 14-20 shows the reference service bandwidth of each service for each user. The MA5821 and MA5822 do not support xDSL access.

Table 14-20 Reference service bandwidth of each service for each user

Service Type	Upstream Bandwidth	Downstream Bandwidth	Bandwidth Description
Internet access service	Determined based on the	Determined based on the service package.	Available bandwidth of Internet access service = Committed bandwidth of the service package - VoIP bandwidth - IPTV bandwidth

Service Type	Upstream Bandwidth	Downstream Bandwidth	Bandwidth Description
	service package.		
VoIP service	200 kbit/s	200 kbit/s	The upstream bandwidth and the downstream bandwidth of VoIP service are symmetrical. The actual bandwidth varies with the coding format. This bandwidth is calculated for two POTS ports.
IPTV service (common program)	/	2.5 Mbit/s per channel	IPTV service mainly occupies the downstream bandwidth. The actual bandwidth depends on the coding format, the picture in picture (PiP) information, 10% of the assured bandwidth for burst traffic, and the number of programs that can be concurrently watched by one user (in the case of multiple STBs). The upstream bandwidth is mainly used for transmitting IGMP packets, which requires a little bandwidth and can be ignored.
IPTV service (high definition program)	/	9.7 Mbit/s per channel	

Note

- It is recommended to configure rate limitation on the BRAS or SR, not on the OLTs or MDUs. If the BRAS does not support rate limit, OLTs can limit rates on service flows using traffic profiles.
- Different service packets of MDUs are distinguished by 802.1p priorities and are mapped to GEM ports based on 802.1p priorities so that packets are transmitted to different GEM ports. Each GEM port (each service) can correspond to a T-CONT or all GEM ports share a T-CONT.
- The sum of the assured bandwidth of all ONTs connected to an OLT PON port and the fixed bandwidth of OMCI management channel should be smaller than the GPON upstream bandwidth. Some bandwidth must be reserved for future service expansion.

15 Layer 3 Features

About This Chapter

This topic describes the network layer (Layer 3) features implemented by the system.

When the SCUN, SCUH or MCUD control board is installed, the SPUF board is used to expand specifications of ARP/ND and routing entries and enhance Layer 3 forwarding capability.

The system supports two Layer 3 forwarding modes, which can be configured by the **router mode** command. Specifications of the Layer 3 feature depend on the control board type and Layer 3 forwarding mode of the system.

- Basic mode: the control board is the Layer 3 forwarding board. This is the default Layer 3 forwarding mode.
- Enhanced mode: the SPUF board is the Layer 3 forwarding board.

For details about functions, specifications, and limitations of the SPUF board, see the related features and Comparison Between SPU Service Processing Boards of the Hardware Description.

15.1 Configuring Layer 3 Forwarding Mode

The enhanced Layer 3 forwarding mode expand specifications of ARP or neighbor discovery (ND) entries and routing entries and enhance Layer 3 forwarding capability.

Prerequisites

When the SCUN, SCUH or MCUD control board is installed, the SPUF board is used.

Context

NOTICE

When the Layer 3 forwarding mode is switched, services and the connection to the NMS are interrupted and the system will reboot. It is recommended that you configure the Layer 3 forwarding mode during a deployment or upgrade and do not configure it during normal service running.

Procedure

Run the **router mode** command to configure the Layer 3 forwarding mode of the system.

- When the Layer 3 forwarding mode is set to **basic**, the control board serves as the Layer 3 forwarding board. This is the default Layer 3 forwarding mode.
- When the Layer 3 forwarding mode is set to **enhanced**, the SPUF board serves as the Layer 3 forwarding board.

Step 1 Run the **display router mode** command to query the Layer 3 forwarding mode of the system.

----End

15.2 ARP

ARP implements conversion between IP addresses and MAC addresses.

15.2.1 Introduction to ARP

Definition

The Address Resolution Protocol (ARP) is a protocol which is used to convert an IP address to a MAC address. It belongs to the TCP/IP protocol suite.

Purpose

The IP address represents only the network layer address of a host. If a host in a network needs to send the network layer data to a destination host, the host must know the physical address (MAC address) of the destination host. Therefore, an IP address has to be translated into a MAC address. ARP is used for translating an IP address to a MAC address.

15.2.2 ARP Principle

ARP Mapping List

Every host has a table named the ARP mapping list for converting IP addresses into MAC addresses.

The ARP mapping list of a host contains a series of mappings between IP addresses and associated MAC addresses of other hosts that have communicated with this host recently.

When a host is started, its ARP mapping list is empty.

Implementation of ARP

ARP enables two hosts in a network to interconnect with each other at Layer 2.

Assume that there are two PCs: host A and host B with IP addresses IP_A and IP_B respectively. Host A sends messages to host B in the following way:

1. Host A checks its ARP mapping list for the ARP mapping entry of IP_B.
   - If host A finds the MAC address of host B, host A encapsulates the IP data packets according to the MAC address and then sends them to host B.
   - If host A does not find the MAC Address of host B, host A puts the data packets in the ARP waiting queue, initiates an ARP request, and then broadcasts it on the Ethernet.

     The ARP request contains the IP address of host B and the IP address and MAC address of host A.
2. As the ARP request is broadcasted, all the hosts on the Ethernet can receive it. Only the requested host (host B), however, responds to the request.
3. Host B stores the IP and MAC addresses of the request initiator (host A) contained in the request, in its own ARP mapping list.
4. Host B returns an ARP response containing the MAC address of host B to host A. Such a response is no longer broadcast, but sent to host A directly.
5. After receiving the response, host A extracts the IP address and MAC address of host B, and adds them to its own ARP mapping list. After that, host A transmits all the data packets in the waiting queue destined for host B.

Static ARP and Dynamic ARP

The manually configured mapping between IP addresses and MAC addresses is known as the static ARP. The mapping between IP addresses and MAC addresses configured dynamically by the ARP protocol is known as the dynamic ARP.

In general, the dynamic ARP is needed. The static ARP is needed only when you need to manually adjust the ARP entries.

A static ARP entry takes effect when the MA5600T/MA5603T/MA5608T works, while the aging time for a dynamic ARP entry is configurable, the default value is 20 minutes.

15.2.3 Configuring ARP Detection (for Accelerating Protection Switching)

Address Resolution Protocol (ARP) probe enables faster protection switching by detecting status of end-to-end links. ARP detection can be configured for a network scenario in which a link protection group is configured for the upstream Ethernet ports on the access device, and there are other types of devices deployed, such as switches and transmission devices, between the access device and the aggregation devices.

Prerequisites

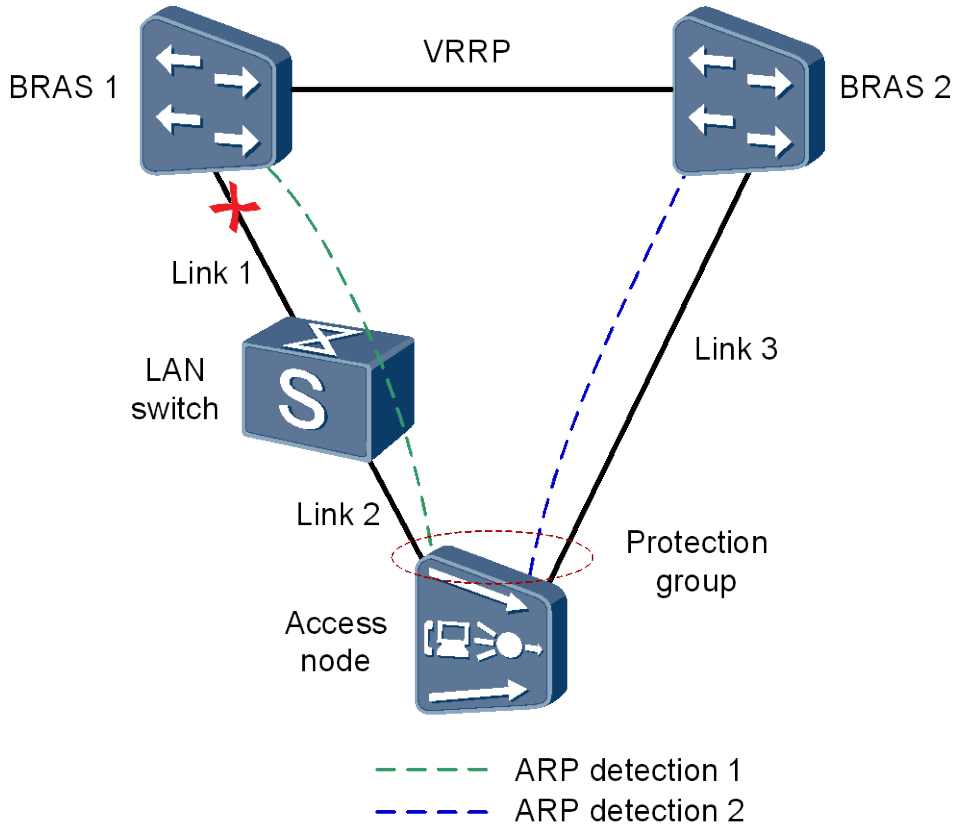
The access device provides upstream ports through the upstream boards. ARP detection is not supported if the access device provides upstream ports through the control board.

A Timedelay protection group is configured on the upstream boards. For configuration details, see 19.4.5 Configuring an Ethernet Port Protection Group.

Context

Figure 15-1 shows an Ethernet port protection example in which the access device (MA5600T/MA5603T/MA5608T) is dual-homed to BRASs.

Figure 15-1 Ethernet port protection with the access device dual-homed to BRASs



The MA5600T/MA5603T/MA5608T is dual-homed to BRAS 1 and BRAS 2. The working links are Link1 and Link2, and the protection link is Link3. If Link1 is faulty and Link2 is normal, the MA5600T/MA5603T/MA5608T can switch services to Link3 by using ARP detection even though the upstream port on the MA5600T/MA5603T/MA5608T is functioning properly. This ensures uninterrupted service transmission.

Procedure

Configure VLAN and Layer 3 port IP address for ARP detection.

1. Create a VLAN.

   Run the **vlan** command to create a smart VLAN for ARP detection.

2. Add an upstream port to the VLAN.

   Run the **port vlan** command to add the working upstream Ethernet port in the protection group to the VLAN. The protection port in the protection group cannot be added to the VLAN.

3. Create a Layer 3 interface for the VLAN.

   Run the **interface vlanif** command to create a Layer 3 interface for the VLAN and enter the VLAN interface mode.

4. Configure an IP address for the Layer 3 interface in the VLAN.

   Run the **ip address** command to configure an IP address for the Layer 3 interface in the VLAN. Ensure that this IP address is in the same subnet as the IP address of the remote device.

Step 1 Configure ARP detection for the working port in the protection group.

1. Configure ARP detection for the working port in the protection group.

   Run the **arp-detect** command to configure ARP detection for the working port in the protection group.

2. (Optional) Configure the times for sending ARP detection packets.

   Run the **detect-multiplier** command to configure the times for sending ARP packets. If the remote device does not respond to the ARP detection packets sent by the local device (here, the access device) for the specified times, the local device considers ARP detection has timed out.

   The waiting time for ARP detection is derived from the following formula: Waiting time = Interval for sending ARP request packets x Times for sending ARP packets. The ARP detection waiting time is also the time taken for triggering a protection switching. The minimum waiting time is 3s (1s x 3). Because the interconnected devices have to process ARP packets, the device CPU load will increase. The more frequent the packets are sent, the heavier the CPU load.

3. (Optional) Configure the interval for sending ARP detection packets.

   Run the **min-tx-interval** command to configure the interval for sending ARP detection packets.

4. Enable the ARP detection function.

   Run the **detect enable** command to enable the ARP detection function.

Step 2 Configure ARP detection for the protection port in the protection group.

Repeat Step 2(but change the working port to the protection port) to configure ARP detection for the protection port in the protection group.

Step 3 Verify ARP detection configurations at the two ports in the protection group.

Run the **display arp-detect** command to verify ARP detection configurations, such as the remote IP address and enable/disable status of ARP detection, at the two ports in the protection group.

----End

Example

Table 15-1 Data plan

Item	Value
Positions of ports in the protection group	Ports on the GIU board: 0/19/0 and 0/19/1
VLAN for ARP detection	VLAN 20

Item	Value
IP address of the Layer 3 interface in the VLAN	Working port: 1.1.1.2/24 Protection port: 2.2.2.2/24
Remote IP address	1.1.1.1/24 2.2.2.1/24
ARP detection times	Three times (default)
Interval for sending ARP detection packets	1s (default)

This example assumes a scenario in which the MA5600T/MA5603T/MA5608T is dual-homed to BRAS 1 and BRAS 2 through the GIU upstream board, and the "Value" column in Table 15-1 lists the data plan. When ARP detection times out, the system considers the working link interrupted and switches services to BRAS 2, ensuring uninterrupted service transmission.

To configure ARP detection in such a network scenario, do as follows:

```
//Configure the VLAN and Layer 3 interface IP address used for ARP detection of the
local device.
huawei(config)#vlan 20 smart
huawei(config)#port vlan 20 0/19 0
huawei(config)#interface vlanif 20
huawei(config-if-vlanif20)#ip address 1.1.1.2 24
huawei(config-if-vlanif20)#ip address 2.2.2.2 24 sub
huawei(config-if-vlanif20)#quit
//Configure ARP detection for the working port.
huawei(config)#arp-detect arp_test1 bind peer-ip 1.1.1.1 vlan 20 port 0/19/0
huawei(config-arp-detect-arp_test1)# detect-multiplier 3
huawei(config-arp-detect-arp_test1)# min-tx-interval 2
huawei(config-arp-detect-arp_test1)#detect enable
huawei(config-arp-detect-arp_test1)#quit
//Configure ARP detection for the protection port.
huawei(config)#arp-detect arp_test2 bind peer-ip 2.2.2.1 vlan 20 port 0/19/1
huawei(config-arp-detect-arp_test2)#detect-multiplier 3
huawei(config-arp-detect-arp_test2)#min-tx-interval 2
huawei(config-arp-detect-arp_test2)#detect enable
huawei(config-arp-detect-arp_test2)#quit
//Query configurations of the working port.
huawei(config)#display arp-detect arp_test1
 ------------------------------------------------------------------------------
  Name      : arp-test2                     Admin State : Enable
  Peerip    : 1.1.1.1                        Interval   : 2(s)
  Vlan      : 20                           Multiplier  : 3
  F/S/P     : 0/19/0                        State      : Down
 ------------------------------------------------------------------------------
//Query configurations of the protection port.
huawei(config)#display arp-detect arp_test2
 ------------------------------------------------------------------------------
  Name      : arp_test2                     Admin State : Enable
  Peerip    : 2.2.2.1                        Interval   : 2(s)
```

```
 Vlan      : 20                            Multiplier  : 3
 F/S/P     : 0/19/1                         State      : Down
-------------------------------------------------------------------------------
```

15.2.4 ARP Reference Standards and Protocols

The following lists the reference documents of ARP:

- IETF RFC 826: An Ethernet Address Resolution Protocol or Converting Network Protocol Addresses to 48-bit Ethernet Address for Transmission on Ethernet Hardware

15.3 ARP Proxy

ARP proxy is a process of handling the ARP requests. It is used to communication between users in different VLAN or in the same VLAN but isolated from each other by layer 3 forwarding.

15.3.1 Introduction to ARP proxy

Definition

When a host sends an ARP request to another host, the request is processed by the access device connected to the two hosts. This process is called ARP proxy.

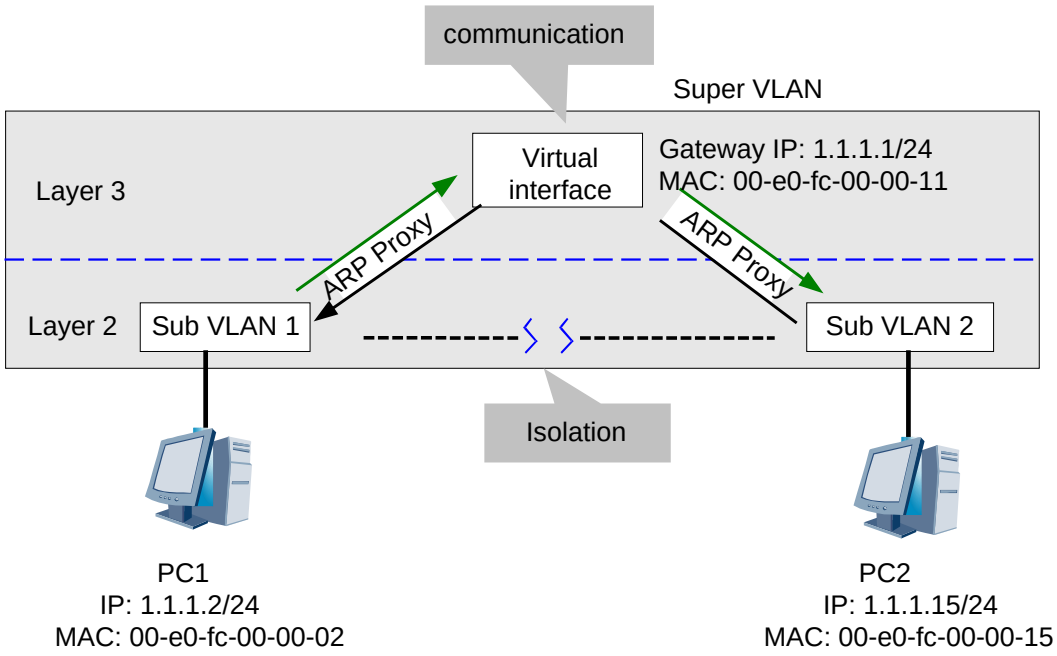
Purpose

On the MA5600T/MA5603T/MA5608T, ARP proxy is often used for interconnection between sub VLANs in a super VLAN.

15.3.2 ARP proxy Principle

As shown in Figure 15-2, PC 1 is in sub VLAN 1, and PC 2 is in sub VLAN 2. They are isolated at Layer 2. PC 1, PC 2 and the virtual Layer 3 interface are in the same subnet.

Figure 15-2 Implementation of the ARP proxy



The following describes how PC 1 and PC 2 communicate with each other.

1. Because PC 1 and PC 2 are in the same subnet, when PC 1 attempts to send packets to PC 2, it broadcasts ARP packets directly to request the MAC address of PC 2.

   Because PC 1 and PC 2 are in different broadcast domains, PC 1 does not receive the ARP response packet from PC 2.
2. When the MA5600T/MA5603T/MA5608T with the ARP proxy enabled receives the ARP request packets, it sends the MAC address of its virtual Layer 3 interface to PC 1, and searches its ARP mapping list for the MAC address of PC 2.
3. If the ARP mapping list contains the MAC address of PC 2, the packets from PC 1 can be forwarded to PC 2 through the virtual Layer 3 interface.
4. If the ARP mapping list does not contain the MAC address of PC 2, the MA5600T/MA5603T/MA5608T broadcasts the ARP request packets through its virtual Layer 3 interface to request the MAC address of PC 2.
5. When the MA5600T/MA5603T/MA5608T receives the ARP response packets from PC 2, the MA5600T/MA5603T/MA5608T adds the MAC address of PC 2 to its ARP mapping list. After this, the implementation of the ARP proxy is complete, and PC 1 and PC 2 communicate with each other through the MA5600T/MA5603T/MA5608T.

15.3.3 Configuring ARP Proxy for Interworking

This topic describes how to configure the Address Resolution Protocol (ARP) proxy of the Layer 3 interface so that users on isolated ports of the same broadcast domain or on ports of different broadcast domains can communicate with each other. To reduce the network load, the ARP request packets are limited in a VLAN.

Context

- By default, the ARP proxy function is disabled.

- When the ARP proxy function is enabled for a super VLAN, sub VLANs in the super VLAN can communicate with each other. To enable the nodes in a sub VLAN to communicate with each other, the ARP proxy function must be enabled for the sub VLAN. Configurations for different scenarios are as follows:
  - In a scenario for Layer 3 communication between users in different VLANs, enable the ARP proxy function for the system, for the super VLAN, and for the sub VLANs in the super VLAN.
  - In a scenario for Layer 3 communication between users in the same VLAN, enable the ARP proxy function for the system, and for the VLAN. The super VLAN does not need to be created.

Networking

Figure 15-3 and Figure 15-4 shows an example network of the ARP proxy.

PC1 and PC2 are in sub VLAN 10, service ports are isolated, and PC3 is in sub VLAN 20. User packets can be forwarded in the Layer 3 forwarding mode through the super VLAN interface. The IP address of the super VLAN interface is 10.0.0.254, and the interface is in the same subnet as PC1, PC2, and PC3. After the ARP proxy function is enabled, PC1 and PC2 can communicate with each other, and PC3 can communicate with PC1 and PC2.

Figure 15-3 Example network of the ARP proxy in a DSLAM network

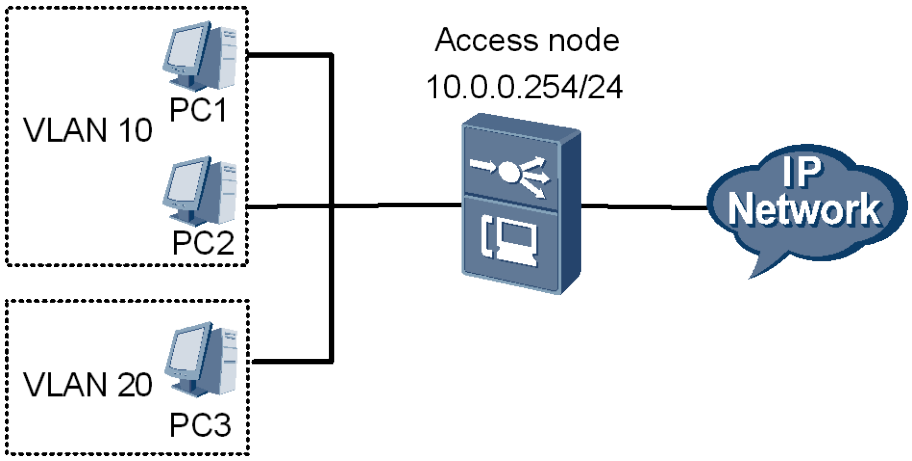
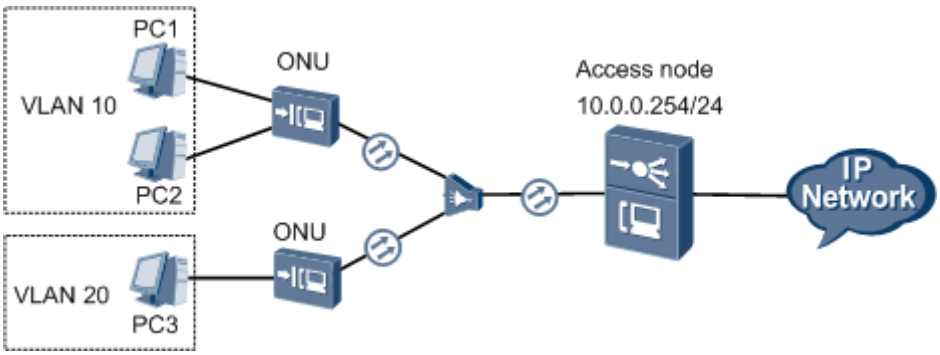


Figure 15-4 Example network of the ARP proxy in an FTTx network



Data Plan

Table 15-2 provides the data plan for configuring the ARP proxy.

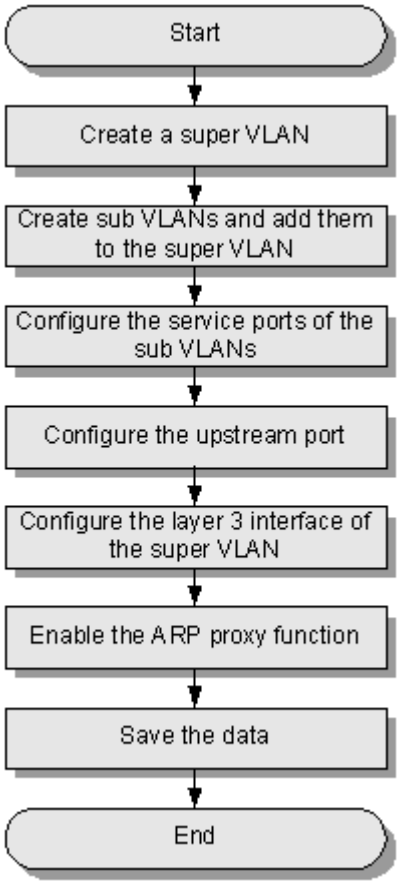
Table 15-2 Data plan for configuring the ARP proxy

Item	Data
Super VLAN	VLAN ID: 100
	Sub VLAN: VLAN 10, VLAN 20
	IP address: 10.0.0.254/24
Sub VLAN	VLAN ID: 10
	VLAN type: smart VLAN
Sub VLAN	VLAN ID: 20
	VLAN type: MUX VLAN
Upstream port	Port: 0/19/0
	VLAN: standard VLAN 30
	IP address: 10.0.1.254/24

Configuration Flowchart

Figure 15-5 shows the flowchart for configuring the ARP proxy.

Figure 15-5 Flowchart for configuring the ARP proxy



Procedure

Create a super VLAN.

```
huawei(config)#vlan 100 super
```

Step 1 Create sub VLANs, and add them to the super VLAN.

```
huawei(config)#vlan 10 smart
huawei(config)#vlan 20 mux
huawei(config)#supervlan 100 subvlan 10
huawei(config)#supervlan 100 subvlan 20
```

Step 2 Configure the service ports of the sub VLANs.

- Configurations in a DSLAM network

```
huawei(config)#service-port vlan 10 adsl 0/2/0 vpi 0 vci 35 rx-cttr 5 tx-cttr 5
huawei(config)#service-port vlan 10 adsl 0/2/1 vpi 0 vci 35 rx-cttr 5 tx-cttr 5
huawei(config)#service-port vlan 20 adsl 0/3/0 vpi 0 vci 35 rx-cttr 5 tx-cttr 5
```

NOTE
VPI/VCI configured on the modem must be 0/35.

- Configurations in an FTTx network

```
huawei(config)#service-port vlan 10 gpon 0/2/0 gemport 128 multi-service user-vlan 15 rx-cttr 5 tx-cttr 5
huawei(config)#service-port vlan 10 gpon 0/2/1 gemport 129 multi-service user-vlan 16 rx-cttr 5 tx-cttr 5
huawei(config)#service-port vlan 20 gpon  0/3/1 gemport 130 multi-service user-vlan 17 rx-cttr 5 tx-cttr 5
```

Step 3 Configure the upstream port.

```
huawei(config)#vlan 30 standard
huawei(config)#port vlan 30 0/19 0
huawei(config)#interface vlanif 30
huawei(config-if-vlanif30)#ip address 10.0.1.254 24
```

NOTE

The IP address of the Layer 3 interface of the super VLAN must be in the same subnet with the IP address obtained by the PC1-PC3.

Step 4 Configure a Layer 3 Interface for the super VLAN

```
huawei(config)#interface vlanif 100
huawei(config-if-vlanif100)#ip address 10.0.0.254 24
```

NOTE

The IP address of the Layer 3 interface of the super VLAN must be in the same subnet with the IP address obtained by the PC.

Step 5 Enable ARP proxy.

1. Enable the ARP proxy function globally.

   ```
   huawei(config)#arp proxy enable
   ```

2. Enable the global ARP proxy on the VLAN interface. After the command is executed, users in different VLANs can communicate with each other.

   ```
   huawei(config-if-vlanif100)#arp proxy enable
   ```

3. Enable ARP proxy on the sub VLAN interface. After the command is executed, users in the same sub VLAN can communicate with each other.

NOTE

Skip this step if you only want PCs in different VLANs to communicate with each other.

```
huawei(config-if-vlanif100)#arp proxy enable subvlan 10
huawei(config-if-vlanif100)#quit
```

Step 6 Save the data.

```
huawei(config)#save
```

----End

Result

- After the global ARP proxy function and the ARP proxy function of the super VLAN interface are enabled, PC1 and PC3, PC2 and PC3 in different VLANs can communicate with each other.

- After the global ARP proxy function, the ARP proxy function of the super VLAN interface, and that of the sub VLAN interface are enabled, PC1 and PC2 in the same VLAN can communicate with each other.

15.3.4 ARP Proxy Reference Standards and Protocols

The following lists the reference documents of ARP proxy:

- IETF RFC1027: Using ARP to Implement Transparent Subnet Gateways

15.4 DHCP Relay

This section describes the implementation and configuration of Dynamic Host Configuration Protocol (DHCP) relay.

15.4.1 What Is DHCP Relay

Definition

DHCP relay enables the MA5600T/MA5603T/MA5608T to forward DHCP packets between DHCP clients and the DHCP server that are in different network segments. DHCP clients can therefore obtain IP addresses dynamically allocated by the DHCP server.

Purpose

The DHCP protocol works in client/server (C/S) mode.

- Multiple DHCP clients request IP addresses from one DHCP server.
- The DHCP server dynamically allocates IP addresses to the DHCP clients.

If the DHCP relay feature is not supported, the DHCP protocol takes effect only if the DHCP clients and the DHCP server are in the same network segment. If they are in different network segments, each network segment requires a DHCP server, which increases deployment costs.

The DHCP relay feature resolves the preceding issue. With this feature, one DHCP server can serve multiple DHCP clients in different network segments. This not only reduces deployment costs but also facilitates centralized management of the DHCP clients.

NOTE

Exchange identification (XID) is a field in a DHCP packet that uniquely identifies the DHCP packet. The MA5600T/MA5603T/MA5608T with DHCP relay enabled changes the XID values of the DHCP packets sent from DHCP clients to values different from the XID values of the DHCP packets received by the DHCP server. The DHCP server generally does not check XID values. Therefore, the XID value change generally does not affect services. However, if carriers add data to the XID field for data checks on the DHCP server, a data check may fail, which would affect services.

15.4.2 DHCPv4 Layer 2 Relay Principles

If the MA5600T/MA5603T/MA5608T located between the DHCPv4 clients and the DHCPv4 server does not support routing, the MA5600T/MA5603T/MA5608T acts only as a bridging device. In this case, the DHCPv4 Layer 2 relay feature enables the MA5600T/MA5603T/MA5608T to transparently transmit DHCPv4 packets and the MA5600T/MA5603T/MA5608T does not need to be configured.

The MA5600T/MA5603T/MA5608T processes DHCPv4 packets only if DHCP Option 82 is enabled on the MA5600T/MA5603T/MA5608T. Specifically, the MA5600T/MA5603T/MA5608T adds or removes Option 82 data from the received DHCPv4 packets. For detailed information about DHCP Option 82, see 27.8.2 DHCP Option 82.

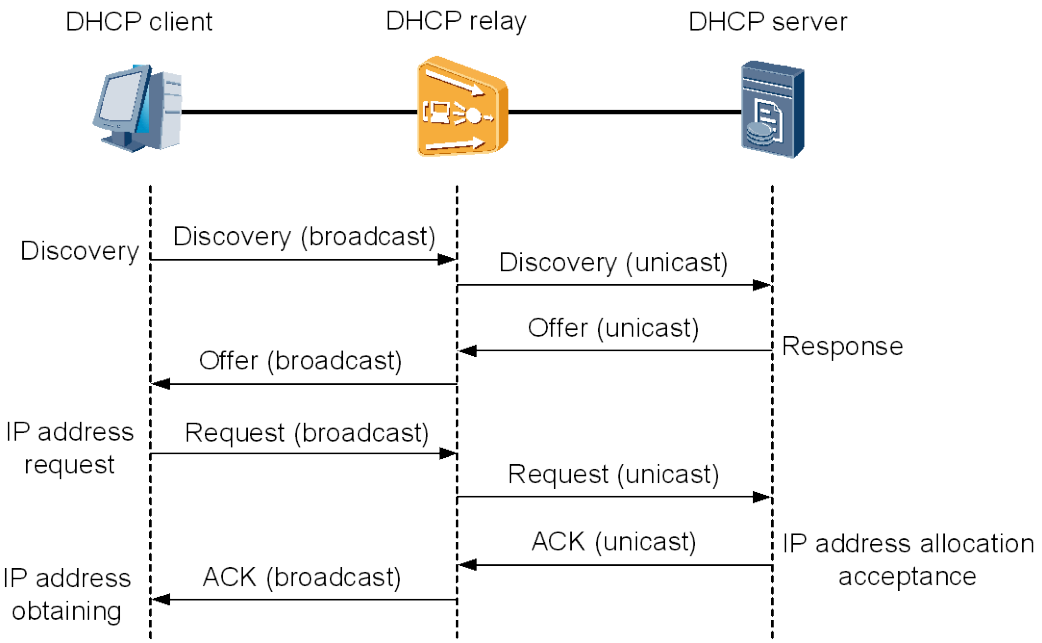
The DHCPv4 Layer 2 relay working process is as follows when DHCP Option 82 is enabled on the MA5600T/MA5603T/MA5608T:

1. A DHCPv4 client broadcasts a request packet during initialization.
2. The MA5600T/MA5603T/MA5608T checks whether the request packet contains Option 82 data.
   - If the request packet contains Option 82 data, the MA5600T/MA5603T/MA5608T retains or replaces the Option 82 data in the request packet according to the configured policy. Then, the MA5600T/MA5603T/MA5608T sends the packet to the DHCPv4 server.
   - If the request packet does not contain Option 82 data, the MA5600T/MA5603T/MA5608T adds Option 82 data to the packet and sends it to the DHCPv4 server. The Option 82 data contains the physical location of the user port initiating the request packet. Therefore, the request packet contains the MAC address of the port on the MA5600T/MA5603T/MA5608T connected to the DHCPv4 client (or the port on the switch connected between the MA5600T/MA5603T/MA5608T and the DHCPv4 client), ID of the VLAN to which the port belongs, and MAC address of the MA5600T/MA5603T/MA5608T.
3. After receiving the DHCPv4 request packet sent from the MA5600T/MA5603T/MA5608T, the DHCPv4 server records the Option 82 data contained in the request packet and sends a packet carrying the DHCPv4 configuration and Option 82 data to the MA5600T/MA5603T/MA5608T.
4. After receiving the packet sent from the DHCPv4 server, the MA5600T/MA5603T/MA5608T removes the Option 82 data from the packet and sends a packet carrying the DHCPv4 configuration to the DHCPv4 client.

15.4.3 DHCPv4 Layer 3 Relay Principles

DHCPv4 Layer 3 relay enables the MA5600T/MA5603T/MA5608T to forward DHCPv4 packets across network segments. DHCPv4 clients can therefore obtain IP addresses dynamically allocated from a DHCPv4 server in a different network segment from the DHCPv4 clients. Figure 15-6 shows the DHCPv4 Layer 3 relay working process.

Figure 15-6 DHCPv4 Layer 3 relay working process



1. During the initialization after a DHCPv4 client starts up, the DHCPv4 client broadcasts a discovery packet in its network segment to search for a DHCPv4 server.
   - If a DHCPv4 server is available in the network segment, the MA5600T/MA5603T/MA5608T is not required and the DHCPv4 client obtains an IP address from the server.
   - If no DHCPv4 server is available in the network segment, the MA5600T/MA5603T/MA5608T is required. The MA5600T/MA5603T/MA5608T receives the discovery packet and unicasts the packet to a DHCPv4 server in another network segment.
2. The DHCPv4 server unicasts an offer packet to the MA5600T/MA5603T/MA5608T to confirm the IP address application. The MA5600T/MA5603T/MA5608T receives the offer packet and broadcasts it to the DHCPv4 client.
3. The DHCPv4 client broadcasts a request packet to request an IP address. The MA5600T/MA5603T/MA5608T receives the request packet and unicasts it to the DHCPv4 server.
4. The DHCPv4 server issues the DHCPv4 configuration to the DHCPv4 client through the MA5600T/MA5603T/MA5608T according to the data carried in the request packet. In this way, the DHCPv4 server dynamically configures the DHCPv4 client.

After DHCP Option 82 is enabled on the MA5600T/MA5603T/MA5608T, the MA5600T/MA5603T/MA5608T adds or removes Option 82 data from received DHCPv4 packets. For details about DHCP packet processing, see Principles.

The MA5600T/MA5603T/MA5608T locates DHCPv4 servers through a DHCPv4 server group. A DHCPv4 server group contains one or more DHCPv4 servers. The MA5600T/MA5603T/MA5608T can select a DHCPv4 server group in one of the following DHCPv4 relay modes.

- Standard mode (default)

This mode is used to specify a DHCPv4 server group for VLAN users. The MA5600T/MA5603T/MA5608T selects a DHCPv4 server group according to the VLAN Layer 3 interface contained in the DHCPv4 packet. The IP address of the VLAN Layer 3 interface is the gateway address of the MA5600T/MA5603T/MA5608T. When using this mode, configure the DHCPv4 server group before binding it to the VLAN Layer 3 interface.

This mode is the simplest among the three DHCPv4 relay modes. In this mode, service types in a VLAN cannot be differentiated.

- Option 60 mode

  This mode is used to specify a DHCPv4 server group for DHCP Option 60 users. The MA5600T/MA5603T/MA5608T selects a DHCPv4 server group according to the Option 60 domain name contained in the DHCPv4 packet. When using this mode, configure the domain name before binding it to the DHCPv4 server group.

  This mode is commonly used. In this mode, service types in a VLAN can be differentiated.

- MAC address segment mode

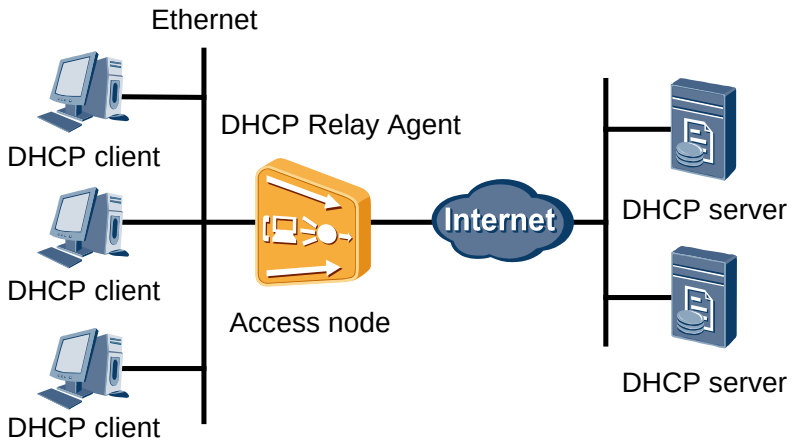
  This mode is used to specify a DHCPv4 server group for the users in a MAC address segment. The MA5600T/MA5603T/MA5608T selects a DHCPv4 server group according to the source MAC address of the DHCPv4 packet. When using this mode, configure the MAC address segment before binding it to the DHCPv4 server group.

  In this mode, service types in a VLAN can be differentiated.

15.4.4 DHCP Relay Networking Applications

Figure 15-7 shows the typical DHCP relay networking.

Figure 15-7 Typical DHCP relay networking



The preceding figure involves the following roles:

- DHCP client: a device that dynamically obtains an IP address or other network configurations from a DHCP server.
- DHCP relay agent (MA5600T/MA5603T/MA5608T): a relay agent for implementing the communication between DHCP clients and a DHCP server in different network segments so that the DHCP clients can obtain IP addresses and other network configurations from the DHCP server. This relieves the need for DHCP server deployment in every network

segment, thereby reducing deployment costs and facilitating centralized DHCP client management.

- DHCP server: a device that allocates IP addresses and other network configurations to DHCP clients.

15.4.5 Configuring DHCP Relay

This section describes how to configure DHCP relay on the MA5600T/MA5603T/MA5608T so that DHCP clients in different network segments from a DHCP server can dynamically obtain IP addresses from the DHCP server.

Context

The MA5600T/MA5603T/MA5608T supports both DHCP Layer 2 relay and DHCP Layer 3 relay. By default, the MA5600T/MA5603T/MA5608T works in DHCP Layer 2 relay mode and transparently transmits DHCP packets. In this mode, the MA5600T/MA5603T/MA5608T does not need to be configured. If the MA5600T/MA5603T/MA5608T works in DHCP Layer 3 relay mode, it requires configuration. The configuration roadmap in the three DHCP relay modes is as follows:

- Standard mode
  a. Configure the DHCP Layer 3 relay mode to **standard**.
  b. Configure a DHCP server group.
  c. Bind the VLAN to the DHCP server group.
- DHCP Option 60 mode
  a. Configure the DHCP Layer 3 relay mode to **option60**.
  b. Configure a DHCP server group.
  c. Create a DHCP Option 60 field.
  d. Bind the DHCP Option 60 field to the DHCP server group.
- MAC address segment mode
  a. Configure the DHCP Layer 3 relay mode to **mac-range**.
  b. Configure a DHCP server group.
  c. Specify a MAC address segment.
  d. Bind the MAC address segment to the DHCP server group.

NOTE
The MA5600T/MA5603T/MA5608T supports DHCP Option 82 to ensure DHCP security. For details about DHCP Option 82, see 27.8.2 DHCP Option 82.

Configuring the Standard Mode

The standard mode is used to specify a DHCP server group for users of a VLAN that contains service ports created on the MA5600T/MA5603T/MA5608T.

Prerequisites

The user VLAN and network VLAN have been configured. For details about how to configure a VLAN, see 13.3.9 Configuring a VLAN.

Procedure

Configure the DHCP forwarding mode.

The DHCP forwarding mode can be configured in global config mode or VLAN service profile mode.

- In global config mode, run the **dhcp mode layer-3 standard** command to configure the DHCP Layer 3 relay mode to **standard**. If you have selected the **vlan** keyword and specified a VLAN ID, the standard DHCP Layer 3 relay mode takes effect only in the specified VLAN.
- In VLAN service profile mode, perform the following operations to configure the DHCP forwarding mode in the user VLAN:
  a. Run the **vlan service-profile** command to create a VLAN service profile.
  b. Run the **dhcp mode layer-3 standard** command to configure the DHCP Layer 3 relay mode to **standard**.
  c. Run the **commit** command to make the profile configuration take effect.
  d. Run the **quit** command to quit the VLAN service profile mode.
  e. Run the **vlan bind service-profile** command to bind the VLAN service profile created in Step 1.a to the VLAN.

Step 1 Configure a DHCP server group.

1. In global config mode, run the **dhcp-server** command to create a DHCP server group.

   igroup-number and *ip-addr* are keywords in the **dhcp-server** command.

   - *igroup-number*: number that uniquely identifies a DHCP server group.

     Before specifying a number for a DHCP server group, check whether the number is unique among the DHCP server group numbers that have been configured on the MA5600T/MA5603T/MA5608T. To do so, run the **display dhcp-server all-group** command on the MA5600T/MA5603T/MA5608T.
   - *ip-addr*: IP address of a DHCP server in a DHCP server group. A maximum of four IP addresses can be specified for this parameter.

   NOTE
   Ensure that the IP address configured in this parameter is the same as the IP address of the network-side DHCP server.

2. (Optional) Run the **dhcp server mode** command to configure the working mode of the DHCP servers in the DHCP server group.

   The DHCP servers in a DHCP server group can work in load sharing or active/standby mode. By default, the DHCP servers work in load sharing mode.

Step 2 Bind the VLAN to the DHCP server.

1. In global config mode, run the **interface vlanif** command to create a VLAN Layer 3 interface.

   Ensure that the value of the **VLANID** keyword in the **interface vlanif** command is the same as the ID of the VLAN that has been created.

2. In VLAN interface (VLANIF) mode, run the **ip address** command to configure the IP address of the VLAN Layer 3 interface.

   After the configuration, all IP packets in the VLAN use this IP address as the source IP address for Layer 3 forwarding.

NOTE

- If all the devices between the MA5600T/MA5603T/MA5608T and the DHCP server work at Layer 2, the IP address of the VLAN Layer 3 interface must be in the same network segment as the IP address of the DHCP server.
- If a device between the MA5600T/MA5603T/MA5608T and the DHCP server works at Layer 3, the IP address of the VLAN Layer 3 interface can be in a network segment different from that of the DHCP server. In this case, ensure that the DHCP server is reachable from the VLAN Layer 3 interface.

3. In VLANIF mode, run the **dhcp-server** command to bind the VLAN to the DHCP server.

   This command requires the *group-number*. Ensure that the *group-number* value is the same as the number of the DHCP server group.

----End

Example

The following is an example of the configurations used to enable DHCP relay on the MA5600T/MA5603T/MA5608T so that DHCP clients in VLAN 2 can obtain IP addresses from DHCP server group 1:

- The DHCP clients in VLAN 2 are in a network segment different from DHCP server group 1.
- The IP address of the Layer 3 interface in VLAN 2 is 10.1.1.101/24.
- DHCP server group 1 contains two DHCP servers working in active/standby mode.
  - The IP address of the active server is 10.1.1.9.
  - The IP address of the standby server is 10.1.1.10.
- The maximum response time of the DHCP servers is 20s.
- The maximum number of DHCP server response timeout times is 10.

```
huawei(config)#dhcp mode layer-3 standard
huawei(config)#dhcp server mode backup 20 10
huawei(config)#dhcp-server 1 ip 10.1.1.9 10.1.1.10
huawei(config)#interface vlanif 2
huawei(config-if-vlanif2)#ip address 10.1.1.101 24
huawei(config-if-vlanif2)#dhcp-server 1
```

Configuring the DHCP Option 60 Mode

The DHCP Option 60 mode is used to specify a DHCP server group for DHCP Option 60 users.

Prerequisites

- The user VLAN and network VLAN have been configured. For details about how to configure a VLAN, see 13.3.9 Configuring a VLAN.
- Before configuring the DHCP Option 60 mode, ensure that the Option 60 domain name is available.

Context

If the MA5600T/MA5603T/MA5608T provisions multiple services, such as the multicast video and VoIP services, the services may be provided by different service providers. These service providers use different DHCP servers or different relay IP addresses of the same DHCP server to allocate IP addresses to DHCP users. Therefore, the DHCP Option 60 mode needs to be configured for the users in different DHCP Option 60 domains to apply for IP addresses.

Procedure

Configure the DHCP forwarding mode.

The DHCP forwarding mode can be configured in global config mode or VLAN service profile mode.

- In global config mode, run the **dhcp mode layer-3 option60** command to configure the DHCP Layer 3 relay mode to **option60**. If you have selected the **vlan** keyword and specified a VLAN ID, the Option 60 DHCP Layer 3 relay mode takes effect only in the specified VLAN.
- In VLAN service profile mode, perform the following operations to configure the DHCP forwarding mode in the user VLAN:
  a. Run the **vlan service-profile** command to create a VLAN service profile.
  b. Run the **dhcp mode layer-3 option60** command to configure the DHCP Layer 3 relay mode to **option60**.
  c. Run the **commit** command to make the profile configuration take effect.
  d. Run the **quit** command to quit the VLAN service profile mode.
  e. Run the **vlan bind service-profile** command to bind the VLAN service profile created in Step 1.a to the VLAN.

Step 1 Configure a DHCP server group.

1. In global config mode, run the **dhcp-server** command to create a DHCP server group.

   igroup-number and *ip-addr* are keywords in the **dhcp-server** command.

   - *igroup-number*: number that uniquely identifies a DHCP server group.

     Before specifying a number for a DHCP server group, check whether the number is unique among the DHCP server group numbers that have been configured on the MA5600T/MA5603T/MA5608T. To do so, run the **display dhcp-server all-group** command on the MA5600T/MA5603T/MA5608T.
   - *ip-addr*: IP address of a DHCP server in a DHCP server group. A maximum of four IP addresses can be specified for this parameter.

   NOTE
   Ensure that the IP address configured in this parameter is the same as the IP address of the network-side DHCP server.

2. (Optional) Run the **dhcp server mode** command to configure the working mode of the DHCP servers in the DHCP server group.

   The DHCP servers in a DHCP server group can work in load sharing or active/standby mode. By default, the DHCP servers work in load sharing mode.

Step 2 In global config mode, run the **dhcp domain** command to create a DHCP Option 60 domain.

Configure the domain name based on the type of the terminal connected to the MA5600T/MA5603T/MA5608T. For example, if the terminal connected to the

MA5600T/MA5603T/MA5608T is a DHCP client running the Windows 98, Windows 2000, Windows XP, or Windows NT OS, the domain name must be **msft**.

Step 3 In DHCP Option 60 mode, run the **dhcp-server** command to bind the DHCP Option 60 domain to the DHCP server group.

After the configuration, the DHCP server group serves all DHCP clients in the DHCP Option 60 domain.

Step 4 Configure the gateway address of the DHCP Option 60 domain.

1. In global config mode, run the **interface vlanif** command to create a VLAN Layer 3 interface.

   Ensure that the value of the **VLANID** keyword in the **interface vlanif** command is the same as the ID of the VLAN that has been created.

2. In VLANIF mode, run the **ip address** command to configure the IP address of the VLAN Layer 3 interface.

   After the configuration, all IP packets in the VLAN use this IP address as the source IP address for Layer 3 forwarding.

NOTE

- If all the devices between the MA5600T/MA5603T/MA5608T and the DHCP server work at Layer 2, the IP address of the VLAN Layer 3 interface must be in the same network segment as the IP address of the DHCP server.
- If a device between the MA5600T/MA5603T/MA5608T and the DHCP server works at Layer 3, the IP address of the VLAN Layer 3 interface can be in a network segment different from that of the DHCP server. In this case, ensure that the DHCP server is reachable from the VLAN Layer 3 interface.

3. In VLANIF mode, run the **dhcp domain gateway** command to configure the gateway address of the DHCP Option 60 domain.

   Ensure that this gateway address is one of the IP addresses of the VLAN Layer 3 interface. Different gateways can be configured for different DHCP Option 60 domains under the same VLAN Layer 3 interface. Therefore, the DHCP Option 60 users connected to different DHCP servers can be differentiated based on their DHCP Option 60 domain name.

----End

Example

The following is an example of the configurations used to enable DHCP relay on the MA5600T/MA5603T/MA5608T so that DHCP clients in VLAN 2 can obtain IP addresses from DHCP server group 1.

- The DHCP clients run the Windows 98, Windows 2000, Windows XP, or Windows NT OS.
- The DHCP clients in VLAN 2 are in a network segment different from DHCP server group 1.
- The IP address of the Layer 3 interface in VLAN 2 is 10.1.2.1/24.
- DHCP server group 1 contains two DHCP servers working in load sharing mode.
  - The IP address of the active server is 10.10.10.10.
  - The IP address of the standby server is 10.10.10.11.

```
huawei(config)#dhcp mode layer-3 Option60
huawei(config)#dhcp-server 1 ip 10.10.10.10 10.10.10.11
```

```
huawei(config)#dhcp domain msft
huawei(config-dhcp-domain-msft)#dhcp-server 1
huawei(config-dhcp-domain-msft)#quit
huawei(config)#interface vlanif 2
huawei(config-if-vlanif2)#ip address 10.1.2.1 24
huawei(config-if-vlanif2)#dhcp domain msft gateway 10.1.2.1
```

Configuring the Gateway Selection Policy in DHCP Option 60 Mode (D-CCAP)

This configuration procedure is used to configure the DHCP relay gateway group and CPE gateway selection policy. The configuration procedure applies to a D-CCAP network where the DHCP server group is selected based on the user's DHCP option 60 domain (also referred to as the DHCP domain in this document).

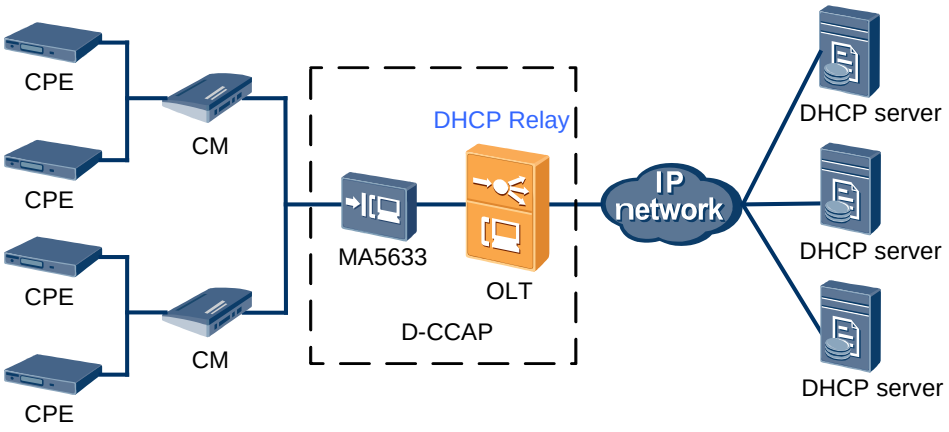
Prerequisites

A VLAN has been created. For details, see 13.3.9 Configuring a VLAN.

Context

As shown in the following figure, in a D-CCAP network, the OLT functions as a DHCP relay, and the CPE dials up using DHCP. The DHCP relay first locates the CM corresponding to the CPE, and then locates the DHCP server through the gateway group to which the CM belongs. The DHCP server allocates IP addresses of different network segments to the CPEs according to the service types.

Figure 15-8 DHCP relay network diagram (D-CCAP)



Procedure

Configure the DHCP forwarding mode.

Configure the DHCP forwarding mode using either of the following methods:

- In global config mode, run the **dhcp mode layer-3 option60** command to set the DHCP relay mode to Layer 3 option60 mode (using parameters **layer-3** and **option60**). If keyword **vlan** is selected and *vlanid* is entered, this configuration takes effect on only this VLAN.

- Perform the configuration in the VLAN service profile:
  a. Run the **vlan service-profile** command to create a VLAN service profile and enter the VLAN service profile mode.
  b. Run the **dhcp mode layer-3 option60** command to configure the DHCP mode.
  c. Run the **commit** command to make the profile configuration take effect. The configuration of the VLAN service profile takes effect only after this command is executed.
  d. Run the **quit** command to quit the VLAN service profile mode.
  e. Run the **vlan bind service-profile** command to bind the VLAN to the VLAN service profile created in Step 1.a.

Step 1 Configure the DHCP server group.

1. In global config mode, run the **dhcp-server** command to create a DHCP server group.
   - *igroup-number*: Indicates the number of the DHCP server group. It identifies a server group. You can run the **display dhcp-server all-group** command to query the DHCP server groups that are already configured and select a DHCP server group number that is not used by the system.
   - *ip-addr*: Indicates the IP address of the DHCP server in the DHCP server group. Up to four IP addresses can be entered.

   NOTE
   The IP address of the DHCP server configured here must be the same as the IP address of the DHCP server on the network side.
2. (Optional) Run the **dhcp server mode** command to configure the working mode of the DHCP server.

   The DHCP servers in the DHCP server group can work in the load-balancing mode or active/standby mode. By default, they work in the load-balancing mode.

Step 2 Configure the DHCP gateway group.

1. In global config mode, run the **dhcp gateway-group** command to create a DHCP gateway group and enter the DHCP gateway group mode.

   A maximum of 64 DHCP gateway groups are supported.
2. In DHCP gateway group mode, run the **dhcp-gateway** command to configure the gateway IP addresses of the DHCP gateway group.
   - Use the **master** parameter to configure the IP address for the master gateway.
   - Use the **slave** parameter to configure the IP addresses for slave gateways.

   A maximum of 1 master gateway and 8 slave gateways can be configured for 1 DHCP gateway group. A slave gateway IP address can be configured only after the master gateway IP address is configured. A master gateway IP address can be deleted only after all slave gateway IP addresses in the same gateway group are deleted.
3. In DHCP gateway group mode, run the **dhcp domain gateway** command to configure the gateway IP addresses corresponding to the DHCP option 60 domain of the CPEs.

   One DHCP gateway group can be configured with 1 default DHCP domain and 4 other DHCP domains. A maximum of 8 gateway IP addresses can be configured for the DHCP gateway group under 1 DHCP domain. The gateway IP address referenced by the DHCP domain must be the slave gateway IP address.
4. In DHCP gateway group mode, run the **dhcp domain server-group** command to configure the DHCP server group corresponding to the DHCP option 60 domain of the CPEs or the CM.

Step 3 Configure the DHCP option 60 domain of the CM.

1. In global config mode, run the **dhcp domain docsis** command to create the DHCP option 60 domain of the CM and enter the DHCP domain mode.

   The **docsis** parameter supports DOCSIS2.0 and DOCSIS3.0.

2. In DHCP domain mode, run the **dhcp gateway learning enable** command to configure the DHCP gateway learning function of the CM.

   The CM can learn the master gateway IP address of a gateway group among the multiple gateway groups corresponding to the DHCP domain.

3. (Optional) In DHCP domain mode, run the **dhcp-server** command to configure the DHCP server group corresponding to the DHCP option 60 domain of the CM.

   If the mapping between the DHCP option 60 domain of the CM and the DHCP server group is configured in both the DHCP gateway group mode and DHCP domain mode, the configuration of the DHCP gateway group mode takes precedence because the configuration is finer-grained than the DHCP-domain-mode configuration. See Step 3.4.

Step 4 Configure the DHCP option 60 domain of the CPE.

1. In global config mode, run the **dhcp domain** command to create the DHCP option 60 domain of the CPE and enter the DHCP domain mode.

   Because a network usually involves terminals of various types and a large number of DHCP option 60 domain names, it is difficult to obtain the DHCP option 60 domain names of all terminals. Therefore, when the terminal types cannot be differentiated, it is recommended to set the DHCP option 60 domain of the CPEs to **default**. The default value supports all terminals. Specify the DHCP option 60 domain name when the terminal types can be differentiated. Below are the domain names of several typical types of terminals.

   - EMTA: The domain name is **pktc**.
   - STB: The domain name is determined by vendors and is set to **default** if it is uncertain.
   - PC: The domain name is **msft** for Windows 98/2000/XP/NT series and is set to **default** when it is uncertain.

2. In DHCP domain mode, run the **dhcp gateway learning enable** command to configure the DHCP gateway learning function of the CPE.

   If a gateway group is configured with multiple slave gateway IP addresses and the CPE has selected multiple slave gateways from the gateway group, the CPE can learn any of the selected slave gateway IP addresses after the DHCP gateway learning function is enabled.

3. (Optional) In DHCP domain mode, run the **dhcp-server** command to configure the DHCP server group corresponding to the DHCP option 60 domain of the CPE.

   If the mapping between the DHCP option 60 domain of the CPE and the DHCP server group is configured in both the DHCP gateway group mode and DHCP domain mode, the configuration of the DHCP gateway group mode takes precedence because the configuration is finer-grained than the DHCP-domain-mode configuration. See Step 3.4.

Step 5 Configure the gateway group corresponding to the DHCP domain of a VLAN Layer 3 interface.

1. In global config mode, run the **interface vlanif** command to create a VLAN Layer 3 interface.

   The VLAN ID must be the same as the ID of the VLAN described in the prerequisite.

2. In VLANIF mode, run the **ip address** command to configure the IP addresses of the VLAN Layer 3 interface.

   The IP addresses of the VLAN Layer 3 interface must include all gateway IP addresses of the gateway group. After the configuration is completed, the IP addresses are used as the source IP addresses for forwarding the IP packets in the VLAN at Layer 3.

   - If only a Layer 2 device exists between the MA5600T/MA5603T/MA5608T and the DHCP server, the IP address of the VLAN Layer 3 interface must be in the same network segment as the IP address of the DHCP server.
   - If the upper-layer device of the MA5600T/MA5603T/MA5608T is a Layer 3 device, the IP address of the VLAN Layer 3 interface and the IP address of the DHCP server can be in different network segments; however, a route must exist between the VLAN Layer 3 interface and the DHCP server.

3. In VLANIF mode, run the **dhcp domain docsis gateway-group** command to configure the gateway group corresponding to the DHCP option 60 domain of the CM under the VLAN Layer 3 interface.

   A maximum of 64 gateway groups can be configured for 1 DHCP domain. A maximum of 4 DHCP domains can be bound to DHCP gateway groups.

4. In VLANIF mode, run the **dhcp domain gateway-policy** command to configure the DHCP gateway selection policy of the CPE.

   - **trace-location-master**: This policy applies to the D-CCAP centralized management scenario. The gateway is selected based on the CM location, and the master gateway IP address is used.
   - **trace-location-slave**: This policy applies to the D-CCAP centralized management scenario. The gateway is selected based on the CM location, and a slave gateway IP address is used.
   - **trace-rid-master**: This policy applies to the D-CCAP standalone NE scenario. The gateway is selected based on the RID specified in the option 82 field of the CPE, and the master gateway IP address is used.
   - **trace-rid-slave**: This policy applies to the D-CCAP standalone NE scenario. The gateway is selected based on the RID specified in the option 82 field of the CPE, and a slave gateway IP address is used.

Step 6 (Optional) Query the configurations of the DHCP gateway group.

- Run the **display dhcp domain** command to query the DHCP domain information, including information about the corresponding DHCP gateway group.
- Run the **display dhcp gateway-group** command to query the DHCP gateway group information.
- Run the **display dhcp interface** command to query the DHCP configurations of a VLAN Layer 3 interface.

----End

Example

Serving as a DHCP relay, the OLT selects DHCP servers for the CM and CPEs through configuration of DHCP gateway groups. The data plan is as follows.

Configuration Item	Data Plan	Description

Configuration Item	Data Plan	Description
Service VLAN of the D-CCAP	10	-
DHCP domains of the CPEs	default	Applicable to all CPEs.
	STB	-
DHCP server group	DHCP server 0: 192.168.1.1, 192.168.2.1	Allocates IP addresses to the CM.
	DHCP server 1: 192.168.1.2, 192.168.2.2	Allocates IP addresses to the CPE with the **default** DHCP domain.
	DHCP server 2: 192.168.1.3, 192.168.2.3	Allocates IP addresses to the CPE with the **STB** DHCP domain.
Gateway groups	huaweigroup1 • Master gateway IP address: 10.1.1.1 • Slave gateway IP addresses: 10.1.2.1, 10.1.3.1, 10.1.4.1, 10.1.5.1	10.1.2.1 is the gateway IP address of the CPE with the **default** DHCP domain, and 10.1.3.1 is the gateway IP address of the CPE with the **STB** DHCP domain.
	huaweigroup2 • Master gateway IP address: 10.1.10.1 • Slave gateway IP addresses: 10.1.20.1, 10.1.30.1, 10.1.40.1, 10.1.50.1	10.1.20.1 is the gateway IP address of the CPE with the **default** DHCP domain, and 10.1.30.1 is the gateway IP address of the CPE with the **STB** DHCP domain.
Gateway groups of the CPEs	huaweigroup1, huaweigroup2	-
DHCP gateway selection policy of the CPEs	trace-rid-slave	Applicable to the D-CCAP standalone NE scenario. In this scenario, the gateway is selected based on the RID specified in option 82 of the CPE, and a slave gateway IP address is used.

```
huawei(config)#vlan 10 smart
huawei(config)#cable service-vlan 10
huawei(config)#interface vlanif 10
huawei(config-if-vlanif10)#ip address 10.1.1.1 255.255.255.0
huawei(config-if-vlanif10)#ip address 10.1.2.1 255.255.255.0 sub
huawei(config-if-vlanif10)#ip address 10.1.3.1 255.255.255.0 sub
huawei(config-if-vlanif10)#ip address 10.1.4.1 255.255.255.0 sub
huawei(config-if-vlanif10)#ip address 10.1.5.1 255.255.255.0 sub
huawei(config-if-vlanif10)#ip address 10.1.10.1 255.255.255.0 sub
```

```
huawei(config-if-vlanif10)#ip address 10.1.20.1 255.255.255.0 sub
huawei(config-if-vlanif10)#ip address 10.1.30.1 255.255.255.0 sub
huawei(config-if-vlanif10)#ip address 10.1.40.1 255.255.255.0 sub
huawei(config-if-vlanif10)#ip address 10.1.50.1 255.255.255.0 sub
huawei(config-if-vlanif10)#quit
huawei(config)#dhcp mode layer-3 option60
//Configure DHCP servers.
huawei(config)#dhcp-server 0 ip 192.168.1.1 192.168.2.1
huawei(config)#dhcp-server 1 ip 192.168.1.2 192.168.2.2
huawei(config)#dhcp-server 2 ip 192.168.1.3 192.168.2.3
//Configure DHCP domains.
huawei(config)#dhcp domain docsis
huawei(config-dhcp-domain-docsis)#dhcp-gateway learning enable
huawei(config-dhcp-domain-docsis)#quit
huawei(config)#dhcp domain default
huawei(config-dhcp-domain-default)#dhcp-gateway learning enable
huawei(config-dhcp-domain-default)#quit
huawei(config)#dhcp domain stb
huawei(config-dhcp-domain-stb)#dhcp-gateway learning enable
huawei(config-dhcp-domain-stb)#quit
//Configure DHCP gateway groups.
huawei(config)#dhcp gateway-group huaweigroup1
huawei(config-dhcp-gateway-group-huaweigroup1)#dhcp-gateway 10.1.1.1 master
huawei(config-dhcp-gateway-group-huaweigroup1)#dhcp-gateway 10.1.2.1 slave
huawei(config-dhcp-gateway-group-huaweigroup1)#dhcp-gateway 10.1.3.1 slave
huawei(config-dhcp-gateway-group-huaweigroup1)#dhcp-gateway 10.1.4.1 slave
huawei(config-dhcp-gateway-group-huaweigroup1)#dhcp-gateway 10.1.5.1 slave
huawei(config-dhcp-gateway-group-huaweigroup1)#dhcp domain docsis dhcp-server 0
huawei(config-dhcp-gateway-group-huaweigroup1)#dhcp domain default dhcp-server 1
huawei(config-dhcp-gateway-group-huaweigroup1)#dhcp domain stb dhcp-server 2
huawei(config-dhcp-gateway-group-huaweigroup1)#dhcp domain default gateway 10.1.2.1
huawei(config-dhcp-gateway-group-huaweigroup1)#dhcp domain stb gateway 10.1.3.1
huawei(config-dhcp-gateway-group-huaweigroup1)#quit
huawei(config)#dhcp gateway-group huaweigroup2
huawei(config-dhcp-gateway-group-huaweigroup2)#dhcp-gateway 10.1.10.1 master
huawei(config-dhcp-gateway-group-huaweigroup2)#dhcp-gateway 10.1.20.1 slave
huawei(config-dhcp-gateway-group-huaweigroup2)#dhcp-gateway 10.1.30.1 slave
huawei(config-dhcp-gateway-group-huaweigroup2)#dhcp-gateway 10.1.40.1 slave
huawei(config-dhcp-gateway-group-huaweigroup2)#dhcp-gateway 10.1.50.1 slave
huawei(config-dhcp-gateway-group-huaweigroup2)#dhcp domain docsis server-group 0
huawei(config-dhcp-gateway-group-huaweigroup2)#dhcp domain default server-group 1
huawei(config-dhcp-gateway-group-huaweigroup2)#dhcp domain stb server-group 2
huawei(config-dhcp-gateway-group-huaweigroup2)#dhcp domain default gateway 10.1.20.1
huawei(config-dhcp-gateway-group-huaweigroup2)#dhcp domain stb gateway 10.1.30.1
huawei(config-dhcp-gateway-group-huaweigroup2)#quit
//Configure the CM gateway groups, and the gateway selection policy of the CPEs under
the VLAN Layer 3 interface.
huawei(config)#interface vlanif 10
huawei(config-if-vlanif10)#dhcp domain docsis gateway-group huaweigroup1
huawei(config-if-vlanif10)#dhcp domain docsis gateway-group huaweigroup2
huawei(config-if-vlanif10)#dhcp domain stb gateway-policy trace-rid-slave
huawei(config-if-vlanif10)#dhcp domain default gateway-policy trace-rid-slave
huawei(config-if-vlanif10)#quit
//Query the configuration results.
huawei(config)#display dhcp gateway-group huaweigroup1
```

```
  DHCP gateway-group name : huaweigroup1
  DHCP gateway          : 10.1.1.1 master
                        10.1.2.1 slave
                        10.1.3.1 slave
                        10.1.4.1 slave
                        10.1.5.1 slave
  DHCP server group      : domain stb server-group 2
                        domain docsis server-group 0
                        domain default server-group 1
  DHCP domain           : Vlanif10 docsis
huawei(config)#display dhcp gateway-group huaweigroup2
  DHCP gateway-group name : huaweigroup2
  DHCP gateway          : 10.1.10.1 master
                        10.1.20.1 slave
                        10.1.30.1 slave
                        10.1.40.1 slave
                        10.1.50.1 slave
  DHCP server group      : domain stb server-group 2
                        domain docsis server-group 0
                        domain default server-group 1
  DHCP domain           : Vlanif10 docsis
```

Configuring the DHCP MAC Address Segment Mode

The DHCP MAC address segment mode is used to specify a DHCP server group for the users in a MAC address segment.

Prerequisites

The user VLAN and network VLAN have been configured. For details about how to configure a VLAN, see 13.3.9 Configuring a VLAN.

Context

The devices on a network may be from different vendors. Each vendor has a consistent MAC address segment and all devices from a vendor use the vendor's MAC address segment. Configure the DHCP MAC address segment mode for the devices to obtain IP addresses from the DHCP server.

The MA5600T/MA5603T/MA5608T selects a DHCP server group based on a MAC address segment. After a MAC address segment is configured, all DHCP clients in the MAC address segment can obtain IP addresses from the DHCP server group.

Procedure

Configure the DHCP forwarding mode.

The DHCP forwarding mode can be configured in global config mode or VLAN service profile mode.

- In global config mode, run the **dhcp mode layer–3 mac-range** command to configure the DHCP Layer 3 relay mode to **mac-range**. If you have selected the **vlan** keyword and specified a VLAN ID, the DHCP MAC address segment mode takes effect only in the specified VLAN.

- In VLAN service profile mode, perform the following operations to configure the DHCP forwarding mode in the user VLAN:
  a. Run the **vlan service-profile** command to create a VLAN service profile.
  b. Run the **dhcp mode layer–3 mac-range** command to configure the DHCP Layer 3 relay mode to **mac-range**.
  c. Run the **commit** command to make the profile configuration take effect.
  d. Run the **quit** command to quit the VLAN service profile mode.
  e. Run the **vlan bind service-profile** command to bind the VLAN service profile created in Step 1.a to the VLAN.

Step 1 Configure a DHCP server group.

1. In global config mode, run the **dhcp-server** command to create a DHCP server group.

   igroup-number and *ip-addr* are keywords in the **dhcp-server** command.

   - *igroup-number*: number that uniquely identifies a DHCP server group.

     Before specifying a number for a DHCP server group, check whether the number is unique among the DHCP server group numbers that have been configured on the MA5600T/MA5603T/MA5608T. To do so, run the **display dhcp-server all-group** command on the MA5600T/MA5603T/MA5608T.
   - *ip-addr*: IP address of a DHCP server in a DHCP server group. A maximum of four IP addresses can be specified for this parameter.

   NOTE
   Ensure that the IP address configured in this parameter is the same as the IP address of the network-side DHCP server.

2. (Optional) Run the **dhcp server mode** command to configure the working mode of the DHCP servers in the DHCP server group.

   The DHCP servers in a DHCP server group can work in load sharing or active/standby mode. By default, the DHCP servers work in load sharing mode.

Step 2 Specify a MAC address segment.

1. In global config mode, run the **dhcp mac-range** command to create a MAC address segment.

   range-name is the MAC address segment name, which is used only for commenting on the MAC address segment.

2. In MAC address segment mode, run the **mac-range mac-address-start to mac-address-end** command to specify the range of the MAC address segment.

Step 3 In MAC address segment mode, run the **dhcp-server** command to bind the MAC address segment to the DHCP server.

Step 4 Configure the gateway address of the MAC address segment.

1. In global config mode, run the **interface vlanif** command to create a VLAN Layer 3 interface.

   Ensure that the value of the **VLANID** keyword in the **interface vlanif** command is the same as the ID of the VLAN that has been created.

2. In VLANIF mode, run the **ip address** command to configure the IP address of the VLAN Layer 3 interface.

   After the configuration, all IP packets in the VLAN use this IP address as the source IP address for Layer 3 forwarding.

NOTE

- If all the devices between the MA5600T/MA5603T/MA5608T and the DHCP server work at Layer 2, the IP address of the VLAN Layer 3 interface must be in the same network segment as the IP address of the DHCP server.
- If a device between the MA5600T/MA5603T/MA5608T and the DHCP server works at Layer 3, the IP address of the VLAN Layer 3 interface can be in a network segment different from that of the DHCP server. In this case, ensure that the DHCP server is reachable from the VLAN Layer 3 interface.

3. In VLANIF mode, run the **dhcp mac-range gateway** command to configure the gateway address of the MAC address segment.

   Ensure that this gateway address is one of the IP addresses of the VLAN Layer 3 interface. Different gateways can be configured for different MAC address segments under the same VLAN Layer 3 interface. Therefore, the users connected to different DHCP servers can be differentiated based on their DHCP MAC address segment.

----End

Example

The following is an example of the configurations used to enable DHCP relay on the MA5600T/MA5603T/MA5608T so that users in the MAC address segment ranging from 0000-0000-0001 to 0000-0000-0100 in VLAN 2 can obtain IP addresses from DHCP server group 1.

- The DHCP clients in VLAN 2 are in a network segment different from DHCP server group 1.
- The IP address of the Layer 3 interface in VLAN 2 is 10.1.2.1/24.
- DHCP server group 1 contains two DHCP servers working in load sharing mode.
  - The IP address of the active server is 10.10.10.10.
  - The IP address of the standby server is 10.10.10.11.

```
huawei(config)#dhcp mode layer-3 mac-range
huawei(config)#dhcp-server 1 ip 10.10.10.10 10.10.10.11
huawei(config)#dhcp mac-range huawei
huawei(config-mac-range-huawei)#mac-range 0000-0000-0001 to 0000-0000-0100
huawei(config-mac-range-huawei)#dhcp-server 1
huawei(config-mac-range-huawei)#quit
huawei(config)#interface vlanif 2
huawei(config-if-vlanif2)#ip address 10.1.2.1 24
huawei(config-if-vlanif2)#dhcp mac-range huawei gateway 10.1.2.1
```

15.4.6 DHCP Relay Standards and Protocols Compliance

The DHCP relay feature complies with the following standard and protocol:

- RFC 2131: Dynamic Host Configuration Protocol
- DHCPv4 Option 82: RFC 3046

15.5 DHCPv6 Relay

DHCPv6 relay functions (in an IPv6 network topology) in a similar way to DHCPv4 relay (in an IPv4 network topology). For details about specifications and principles, see 15.4 DHCP Relay. This topic describes the differences between DHCPv6 relay and DHCPv4 relay regarding their functions.

15.5.1 DHCPv6 Relay Principle

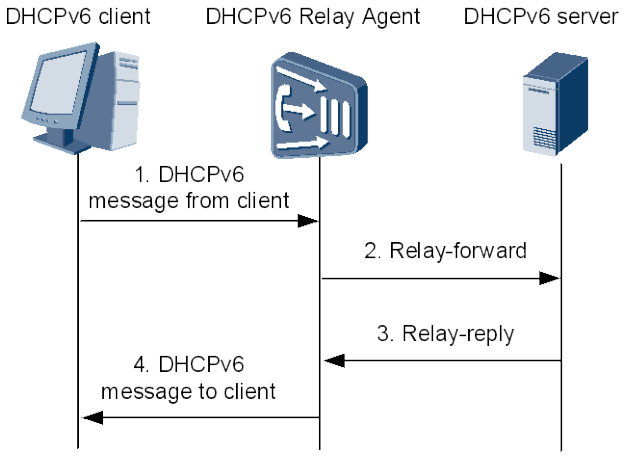
Principle of DHCPv6 L2 Relay

1. When a DHCPv6 relay agent receives a message that needs to be processed using L2 relay, the relay agent constructs a new Relay-forward message, regardless of the original message type. The relay agent copies the IP address of the original message to the peer-address field of the new Relay-forward message, and copies the entire original DHCP message (excluding the IP header and UDP header) to the Relay Message Option (Option 9) of the new Relay-forward message.
2. If the original message originates from a DHCPv6 client, the DHCPv6 relay agent initializes the hop-count field of the Relay-forward message to 0. If the original message is a Relay-forward message that originates from another DHCPv6 relay agent, the relay agent first checks the hop-count field of the message. If the value of the field is greater than or equal to 32, the relay agent discards the message; if the value is smaller than 32, the relay agent adds 1 to the value and uses the new value as the value of the hop-count field of the new Relay-forward message.
3. The DHCPv6 relay agent invariably sets the link-address Field of the Relay-forward message to an unspecified address (::), and invariably includes the interface-id field (Option 18) and remote-id field (Option 37) in the Relay-forward message.
4. In the downstream direction, the Relay Message Option (Option 9) of the original Relay-reply message is extracted, and the content of the Relay Message Option (Option 9) is forwarded as a new downstream message to the DHCPv6 client.

Principle of DHCPv6 L3 Relay

The processing flows between the DHCPv6 client and server are the same when the IPv6 address, IPv6 prefix and other network configuration parameters are dynamically obtained using or not using the DHCPv6 relay agent. Figure 15-9 illustrates only the forwarding process when DHCPv6 relay is used.

Figure 15-9 Working process with DHCPv6 relay



1. The DHCPv6 client sends a request message to the multicast address FF02::1:2 of all DHCPv6 servers and relay agents.
2. After receiving the request message, the DHCPv6 relay agent encapsulates the message into the Relay Message Option of the Relay-forward message and sends the Relay-forward message to the DHCPv6 server.
3. The DHCPv6 server parses the Relay-forward message and obtains the client's request message. Then, the server selects an IPv6 address and other parameters for the client, constructs a reply message, encapsulates the reply message into the Relay Message Option of the Relay-reply message, and sends the Relay-reply message to the DHCPv6 relay agent.
4. The DHCPv6 relay agent parses the Relay-reply message, obtains the server's reply message, and then forwards the message to the DHCPv6 client.
5. The DHCPv6 client implements network configuration according to the IPv6 address, IPv6 prefix, and other parameters contained in the reply message.

NOTE

A DHCPv6 server allocates both IPv6 addresses and IPv6 prefixes to DHCPv6 clients. After obtaining an IPv6 prefix allocated by the DHCPv6 server, a DHCPv6 client sends a Router Advertisement (RA) message containing the IPv6 prefix to the network in which it is located. Using the prefix, hosts in the network can automatically configure their IPv6 addresses.

15.5.2 Differences Between DHCPv4 and DHCPv6 Configurations

Dynamic Host Control Protocol version 6 (DHCPv6) is a DHCP protocol for IPv6. This topic describes the differences between DHCPv4 and DHCPv6 regarding function specifications and commands. Before configuring DHCPv6 services, it is recommended that you familiarize yourself with the procedures and principles of configuring DHCPv4 services and be aware of the differences between DHCPv4 and DHCPv6 configurations.

Context

The differences between DHCPv4 and DHCPv6 configurations are as follows:

- When working in Layer 3 forwarding mode, DHCPv6 supports both the standard mode and option 16 mode. You can run the **dhcpv6 mode { layer-2 | layer-3 [option16] }** command to configure the DHCPv6 working mode. DHCPv4 supports the standard, media access control (MAC) address segment, and DHCP option 60 modes. You can run

the **dhcp mode { layer-2 | layer-3 { mac-range | option60 | standard } }** command to configure the DHCPv4 working mode.

- DHCPv6 does not support DHCP proxy.
- Regarding commands, DHCPv4 uses **dhcp** while DHCPv6 uses **dhcpv6** as the command word. For example, the **dhcp-server** command is used to configure DHCPv4 server groups while the **dhcpv6-server** command is used to configure DHCPv6 server groups.

For the differences of other configuration commands, see "DHCPv6 Configuration" in the *Command Reference*. Some commands need to be executed in diagnose mode. For details about these commands, see "Diagnose Mode Command" in the *Command Reference*.

15.5.3 DHCPv6 Relay Reference Standards and Protocols

The following lists the reference standards and protocols of the DHCPv6 relay feature:

- DHCPv6 relay: RFC 3315 (DHCPv6 protocol created by IETF)
- DHCPv6 Layer 2 relay: draft-ietf-dhc-dhcpv6-ldra-02 (Draft of a DHCPv6 Layer 2 relay protocol created by IETF)
- DHCPv6 option 37: RFC 4649

15.6 DHCP Proxy

This section describes the implementation and configuration of Dynamic Host Configuration Protocol (DHCP) proxy.

15.6.1 What Is DHCP Proxy

Definition

DHCP proxy enables the MA5600T/MA5603T/MA5608T to modify the DHCP packets exchanged between a DHCP server and DHCP clients based on site requirements.

DHCP proxy is comprised of the server ID proxy and lease proxy functions.

- Server ID proxy

  The Option 54 field in a DHCP packet specifies the IP address of a DHCP server. Server ID proxy enables the MA5600T/MA5603T/MA5608T to change the value of the Option 54 field so that the actual IP address of the DHCP server is hidden from DHCP clients. This reduces the risk of DHCP server attacks initiated by DHCP clients.

- Lease proxy

  The IP address lease that a DHCP client applies for is determined based on Options 51, 58, and 59 in a DHCP packet initiated by the DHCP client. Lease proxy enables the MA5600T/MA5603T/MA5608T to change the values of these options so that the MA5600T/MA5603T/MA5608T provides a shorter lease than that provided by the DHCP server for the DHCP client. This facilitates lease management.

Purpose

The DHCP proxy functions offer solutions for different problems.

- Server ID proxy

The actual IP address of the DHCP server can be hidden from DHCP clients, which reduces the risk of DHCP server attacks initiated by DHCP clients.

- Lease proxy

  The duration of the lease configured by the DHCP server is so long that the MA5600T/MA5603T/MA5608T cannot obtain the DHCP client status in a timely manner. This obstructs service provisioning.

  The MA5600T/MA5603T/MA5608T with DHCP proxy enabled obtains the DHCP client status in a timely manner. In addition, the MA5600T/MA5603T/MA5608T can process request packets of the DHCP clients for re-leasing IP addresses without forwarding these packets to the DHCP server for processing. This decreases the load of the DHCP server.

NOTE

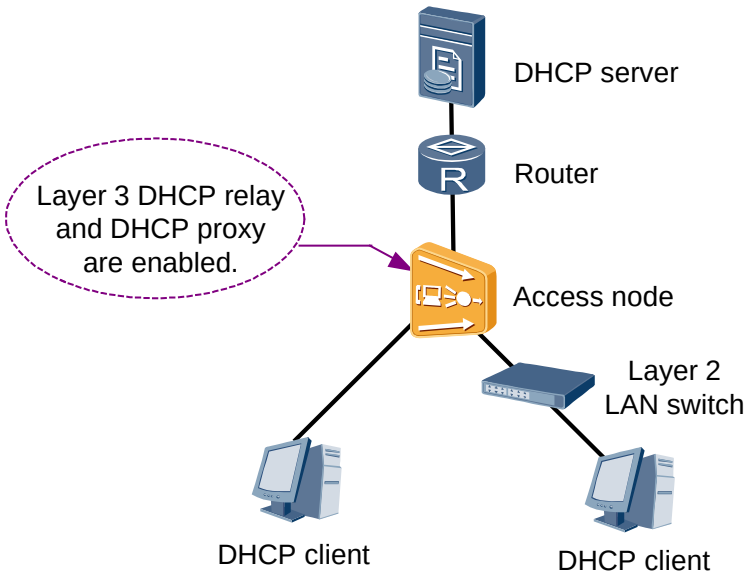
Exchange identification (XID) is a field in a DHCP packet that uniquely identifies the DHCP packet. The MA5600T/MA5603T/MA5608T with DHCP proxy enabled changes the XID values of the DHCP packets sent from DHCP clients to values different from the XID values of the DHCP packets received by the DHCP server. The DHCP server generally does not check XID values. Therefore, the XID value change generally does not affect services. However, if carriers add data to the XID field for data checks on the DHCP server, a data check may fail, which would affect services.

15.6.2 DHCP Proxy Principles

Application Scenario

The MA5600T/MA5603T/MA5608T supports DHCP proxy only when Layer 3 DHCP relay is enabled on it. Both user ports and cascading ports on the MA5600T/MA5603T/MA5608T support DHCP proxy. The MA5600T/MA5603T/MA5608T with DHCP proxy enabled can monitor all DHCP packets exchanged between DHCP clients and a DHCP server. Figure 15-10 shows an application scenario of DHCP proxy.

Figure 15-10 Application scenario of DHCP proxy



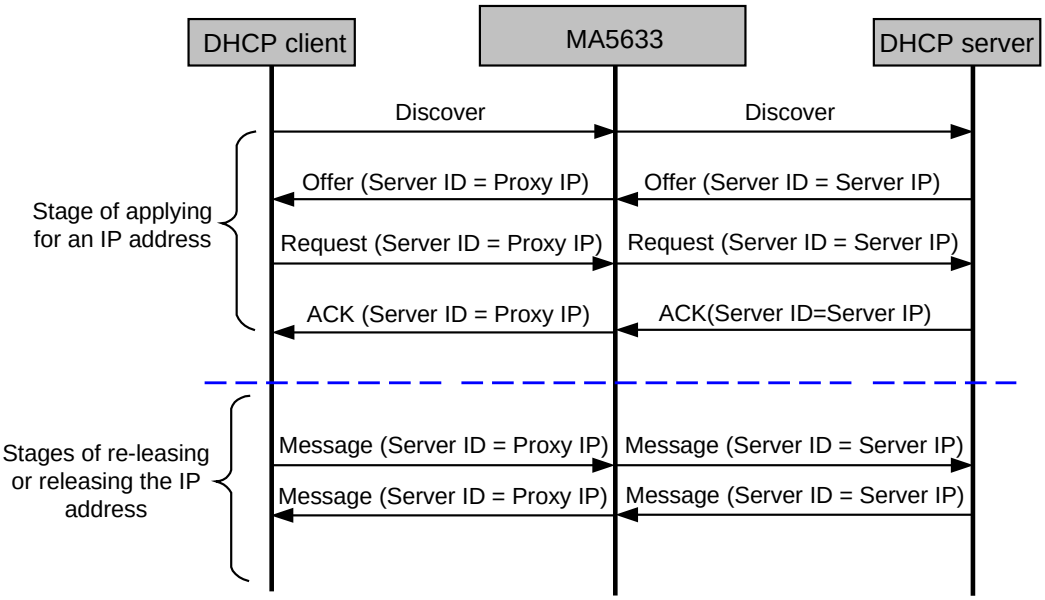
Server ID Proxy

With DHCP proxy enabled, the server ID proxy function alters the exchange of packets as follows:

- In the downstream direction, the MA5600T/MA5603T/MA5608T changes the value of the Option 54 field in the response packets, such as the Offer and ACK packets, sent by the DHCP server to the IP address of the MA5600T/MA5603T/MA5608T. Then, the DHCP client receiving the packets obtains the IP address of the MA5600T/MA5603T/MA5608T instead of the DHCP server's IP address.
- In the upstream direction, the MA5600T/MA5603T/MA5608T restores the value of the Option 54 field in the DHCP packets sent from the DHCP client to the IP address of the DHCP server.

Figure 15-11 shows the exchange of packets between a DHCP client and DHCP server.

Figure 15-11 Exchange of packets between a DHCP client and DHCP server (server ID proxy)



Lease Proxy

Stage of applying for an IP address:

1. The DHCP client sends a request packet to the DHCP server for an IP address. The DHCP server then sends a response packet and allocates an IP address with lease L1 to the DHCP client.
2. The MA5600T/MA5603T/MA5608T captures the response packet from the DHCP server, changes the L1 value in the packet to a smaller value L2 (configurable on the MA5600T/MA5603T/MA5608T), and sends the Offer packet containing lease L2 to the DHCP client. Then, the lease of the IP address allocated to the DHCP client is changed to L2.

Stage of re-leasing the IP address:

1. The DHCP client sends a request packet to the DHCP server to re-lease the IP address when half of the lease's duration has elapsed.

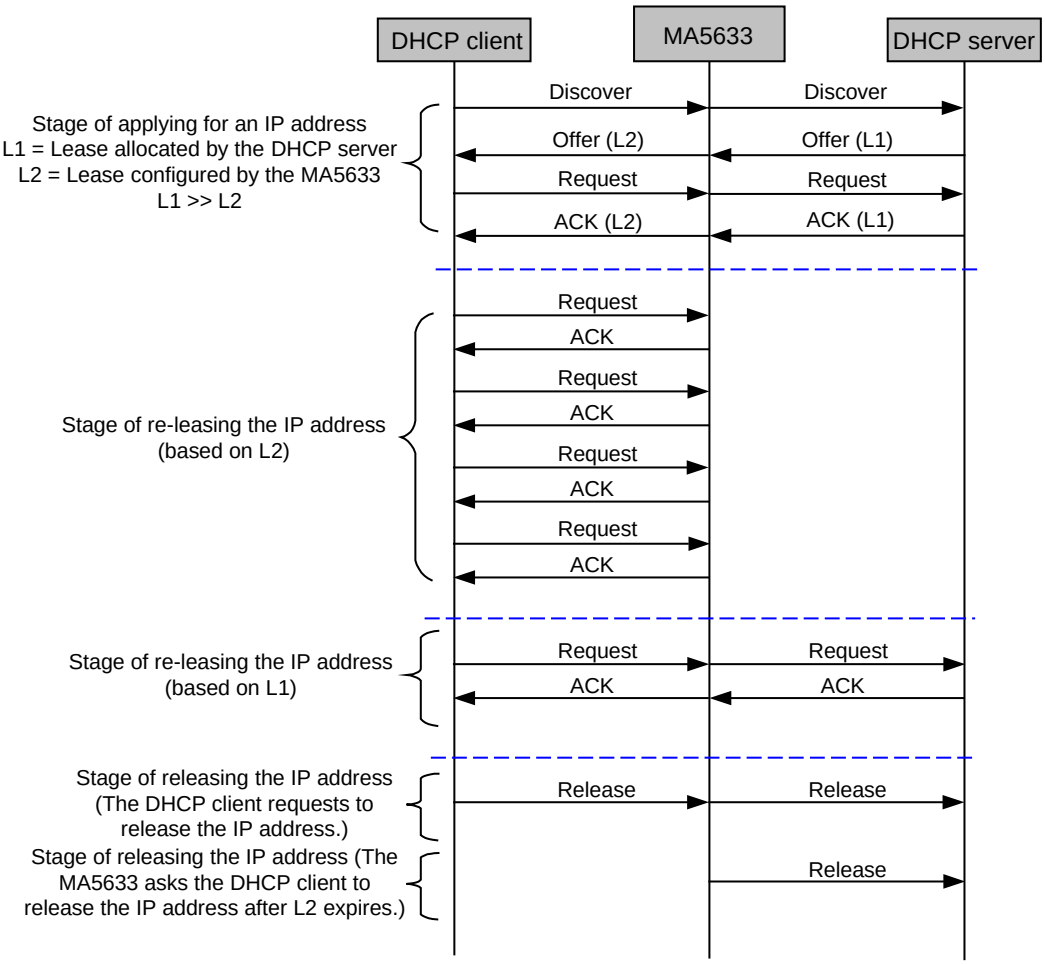
2. The MA5600T/MA5603T/MA5608T captures the request packet and determines whether to forward the request packet to the DHCP server.
   - If the elapsed lease is less than half of L1, the MA5600T/MA5603T/MA5608T responds to the request packet, but does not forward it to the DHCP server. In addition, the MA5600T/MA5603T/MA5608T changes the lease duration to L2.
   - If the elapsed lease is greater than or equal to half of L1, the MA5600T/MA5603T/MA5608T forwards the request packet to the DHCP server.
3. After receiving the request packet, the DHCP server performs the following operations:
   - Responds to the request packet if it allows the re-leasing. Then, the MA5600T/MA5603T/MA5608T forwards the response packet to the DHCP client.
   - Sends a NACK packet to the DHCP client if it does not allow the re-leasing. Then, the MA5600T/MA5603T/MA5608T notifies the DHCP client that this IP address will be released.

Stage of releasing the IP address:

- If the DHCP client sends a request packet to release the IP address, the MA5600T/MA5603T/MA5608T forwards the request packet to the DHCP server.
- If the MA5600T/MA5603T/MA5608T detects that lease L2 of the DHCP client has expired and the MA5600T/MA5603T/MA5608T has not received any request to re-lease the IP address from the DHCP client, the MA5600T/MA5603T/MA5608T sends a request to release the IP address to the DHCP server.

Figure 15-12 shows the exchange of packets between a DHCP client and DHCP server.

Figure 15-12 Exchange of packets between a DHCP client and DHCP server (lease proxy)



15.6.3 DHCP Proxy Standards and Protocols Compliance

The DHCP proxy feature complies with the dsl2006[1].127.00 standard (proposals of DHCP relay improvements).

15.7 VRRP Snooping

VRRP is a fault-tolerant protocol. It allows multiple routers to form a virtual routing device, and provides a mechanism, which ensures that services will be taken over in time by another device once the next hop of a host fails. In this way the continuity and reliability of communication are ensured. VRRP snooping is to snoop (or listen for) VRRP packets.

15.7.1 Introduction to VRRP Snooping

Definition

Virtual Router Redundancy Protocol (VRRP) is a fault-tolerant protocol. It allows multiple routers to form a virtual routing device, and provides a mechanism, which ensures that

services will be taken over in time by another device once the next hop of a host fails. In this way the continuity and reliability of communication are ensured.

VRRP snooping is to snoop (or listen for) VRRP packets. According to VRRP packets the listening device can confirm the port to which the upstream master router is connected. Then, the listening device will transmit the unicast service stream to the master router and at the same time transparently transmit the VRRP packets of any of other routers to another router in the same VRRP group.

Purpose

To enhance system reliability, the MA5600T/MA5603T/MA5608T is directly dual-homed to two or more BRASs in the upstream direction, and the BRASs run the VRRP protocol. When the MA5600T/MA5603T/MA5608T works in the SVLAN+CVLAN forwarding mode and MAC address learning is disabled, the upstream ports of the MA5600T/MA5603T/MA5608T need to be isolated from each other in order to prevent unknown unicast broadcast storm. However, when the upstream ports are isolated, the upstream BRASs cannot interoperate VRRP packets. VRRP snooping is adopted for forwarding VRRP packets because VRRP snooping enables the BRASs to interoperate VRRP packets so that the BRASs can run VRRP normally.

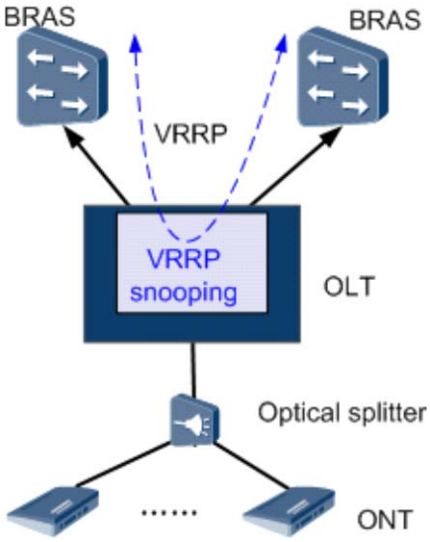
To prevent asynchronous VRRP status, Huawei develops the VRRP Group Management Protocol (VGMP), which is extended based on VRRP. VGMP manages the VRRP status of each backup group in a unified manner. VGMP provides a mechanism for managing the status synchronization, preemption, and channels of multiple VRRP backup groups. When the MA5600T/MA5603T/MA5608T works in the SVLAN+CVLAN mode, the VGMP protocol may fail to run if the upstream ports are isolated. To enable the BRASs to interoperate VGMP packets, the MA5600T/MA5603T/MA5608T can transparently transmit VGMP packets.

15.7.2 VRRP Snooping Principle

Basic Principle of VRRP Snooping

In a network, the failure of a single router may cause failure of the entire network. To address this problem, multiple VRRP-running routers on the upper layer of the MA5600T/MA5603T/MA5608T can form a set of virtual routers. Thus, another router will automatically take over the forwarding service if the master router fails. Viewed from the MA5600T/MA5603T/MA5608T there is still only one router, but this router may be interconnected to two upstream ports of the MA5600T/MA5603T/MA5608T. When the path switches from one port to the other, the upstream router will send free ARP packets to update the forwarding path.

Figure 15-13 Network topology of VRRP snooping



As shown in Figure 15-13, to solve the problem of MAC address insufficiency, the MA5600T/MA5603T/MA5608T adopts the SVLAN+CVLAN forwarding mode. In this forwarding mode, MAC address learning needs to be disabled, and the upstream ports of the MA5600T/MA5603T/MA5608T need to be isolated to avoid broadcast storm of unknown unicast packets.

After the upstream ports of the MA5600T/MA5603T/MA5608T are isolated, the multiple upstream routers directly connected to the MA5600T/MA5603T/MA5608T cannot forward VRRP packets to each other through the upstream port of the MA5600T/MA5603T/MA5608T. The result will be that the routers fail to run the VRRP protocol. To address this problem, the MA5600T/MA5603T/MA5608T needs to employ software forwarding in order to implement VRRP protocol packet exchange between the isolated ports.

When MAC address learning is disabled, packets going upstream may be forwarded to the two upstream ports at the same time, which is a waste of bandwidth. In this case, a static MAC address needs to be configured so that unicast packets are forwarded to the master router only. The MA5600T/MA5603T/MA5608T listens to VRRP packets and free ARP packets to learn the upstream port to which the master router is currently connected. By using the static MAC address, the MA5600T/MA5603T/MA5608T forwards Layer 2 service data to this upstream port.

When the router sends free ARP packets to the MA5600T/MA5603T/MA5608T for switching the forwarding path, the free ARP packets may be lost due to network reasons. If ARP packets are lost, the MA5600T/MA5603T/MA5608T listens to VRRP packets to update the ARP entry. This prevents a condition where Layer 3 forwarding services are interrupted for a long time because the forwarding path is not updated in time.

VRRP Snooping in the VLAN+MAC Forwarding Mode

In the VLAN+MAC forwarding mode, the two routers to which the MA5600T/MA5603T/MA5608T is dual-homed can run the VRRP protocol without additional processing on the MA5600T/MA5603T/MA5608T as long as the following condition is met: The two upstream ports connected to the two routers can interoperate, which allows for normal forwarding of VRRP packets between the two routers. When the network condition is good, or when the MA5600T/MA5603T/MA5608T needs not consider the loss of free ARP

packets (a router may provide for retransmission of free ARP packets), VRRP snooping needs not be enabled on the MA5600T/MA5603T/MA5608T when the MA5600T/MA5603T/MA5608T runs in the VLAN+MAC forwarding mode.

15.7.3 Configuring VRRP Transparent Transmission in the S+C Forwarding Mode

In the S+C forwarding mode, after VRRP snooping is enabled, VRRP packets can be forwarded between two isolated upstream ports.

Context

To enhance the system reliability, the MA5600T/MA5603T/MA5608T is directly connected to two or more routers in the upstream direction for dual homing.

- In the S+C forwarding mode, the MAC address learning function needs to be disabled. In addition, to prevent broadcast storms, two upstream ports of the MA5600T/MA5603T/MA5608T need to be isolated. Therefore, the VRRP packets between upper-layer routers cannot be forwarded through the upstream ports of the MA5600T/MA5603T/MA5608T. To solve this problem, enable the VRRP transparent transmission function.
- In the VLAN+MAC forwarding mode, the VRRP transparent transmission function may not be enabled.

Procedure

Configuring an isolation group.

Run the **isolate group** command to configure an isolation group. To avoid forwarding downstream service packets to another upstream port, you can add the upstream ports to an isolation group.

Step 1 Configure a snooping port.

Run the **vrrp-snoop port** command to configure the upstream port connecting the MA5600T/MA5603T/MA5608T and the router as a snooping port.

Step 2 Configure the virtual IP address and VLAN to be snooped.

Run the **vrrp-snoop ip** command to configure the IP address and VLAN of the virtual router to be snooped.

Step 3 Enable VRRP snooping.

Run the **vrrp-snoop enable** command to enable VRRP snooping.

----End

Example

Assume the following configurations: The MA5600T/MA5603T/MA5608T is connected to two routers through upstream ports 0/19/0 and 0/19/1. The VRRP packets between the routers need to be transparent transmitted through the upstream ports of the MA5600T/MA5603T/MA5608T.The VLAN forwarding mode is S+C, the VLAN ID is 100, and the IP address of the virtual router is 10.71.10.1. To perform these configurations, do as follows:

```
huawei(config)#isolate group port 0/19/0 0/19/1
huawei(config)#vrrp-snoop port 0/19/0
huawei(config)#vrrp-snoop port 0/19/1
huawei(config)#vrrp-snoop ip 10.71.10.1 vlan 100
huawei(config)#vrrp-snoop enable
```

15.7.4 VRRP Snooping Reference Standards and Protocols

The reference standards and protocols of this feature are as follows:

- RFC3768, Virtual Router Redundancy Protocol (VRRP)
- RFC2787, Definitions of Managed Objects for the Virtual Router Redundancy Protocol

15.8 IP-aware Bridge

IP-aware bridge is a Layer 3 forwarding technique. It helps forward data in the upstream direction to different upper-layer devices based on destination IP addresses and configured static IP routes.

15.8.1 Introduction to IP-aware Bridge

Definition

IP-aware bridge is a feature in which an access node can implement Layer 3 forwarding without being configured with an IP address.

Purpose

- To implement Layer 3 forwarding. In this feature, a large number of user MAC addresses can be replaced with the system MAC address of a device for packet forwarding.
- To identify the destination IP address (IP-aware) of users' traffic streams, and send the traffic streams to the corresponding next hop (traffic split) according to route information.
- To terminate user-side ARP requests, terminate network-side ARP requests, and respond by using ARP proxy.
- To implement ARP proxy between users so that users who are in the same VLAN and isolated at Layer 2 can interoperate at Layer 3.

Benefits to Users

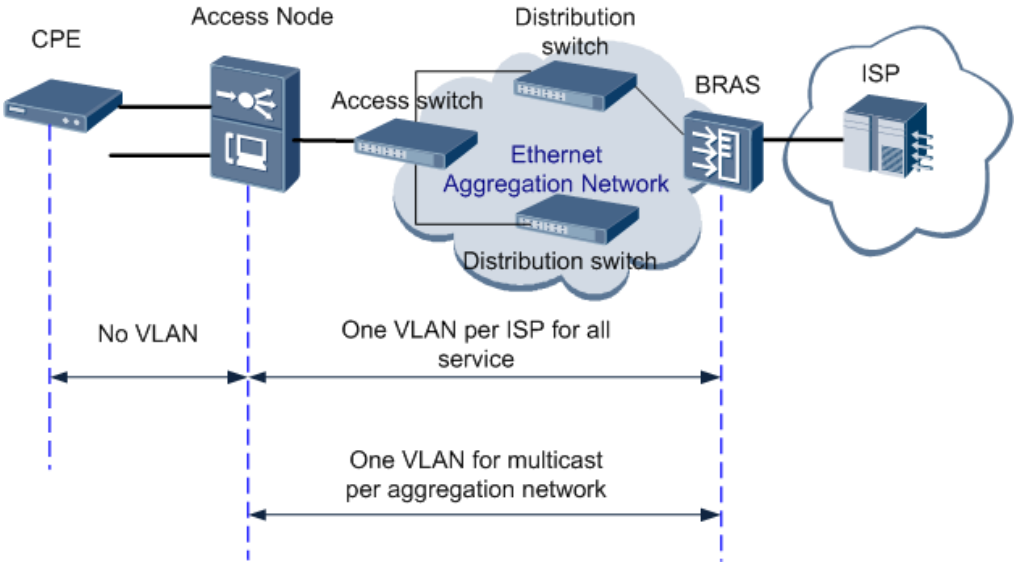
Layer 3 forwarding can be implemented without occupying IP addresses or requiring the configuration of IP addresses.

15.8.2 IP-aware Bridge Principle

Application Scenario

Figure 15-14 shows the application scenario of IP-aware bridge.

Figure 15-14 Application scenario of IP-aware bridge



The DSLAM must meet the following requirements:

- Converts the user MAC address sent by the CPE into the system MAC address.
- Sends the traffic streams of users to different ISPs according to the destination IP addresses.
- Terminates ARP requests:
  - Terminates the ARP requests of users and responds using the system MAC address.
  - Terminates the ARP requests of upper-layer devices to users and responds using the system MAC address.
- Does not need equipment IP address of its own. The DSLAM can query the IP address through the CPE.

Principle Description

Figure 15-15 shows the flow of Layer 3 forwarding of IP-aware bridge in the upstream direction, and Figure 15-16 shows that in the downstream direction.

Figure 15-15 Flow of Layer 3 forwarding in the upstream direction

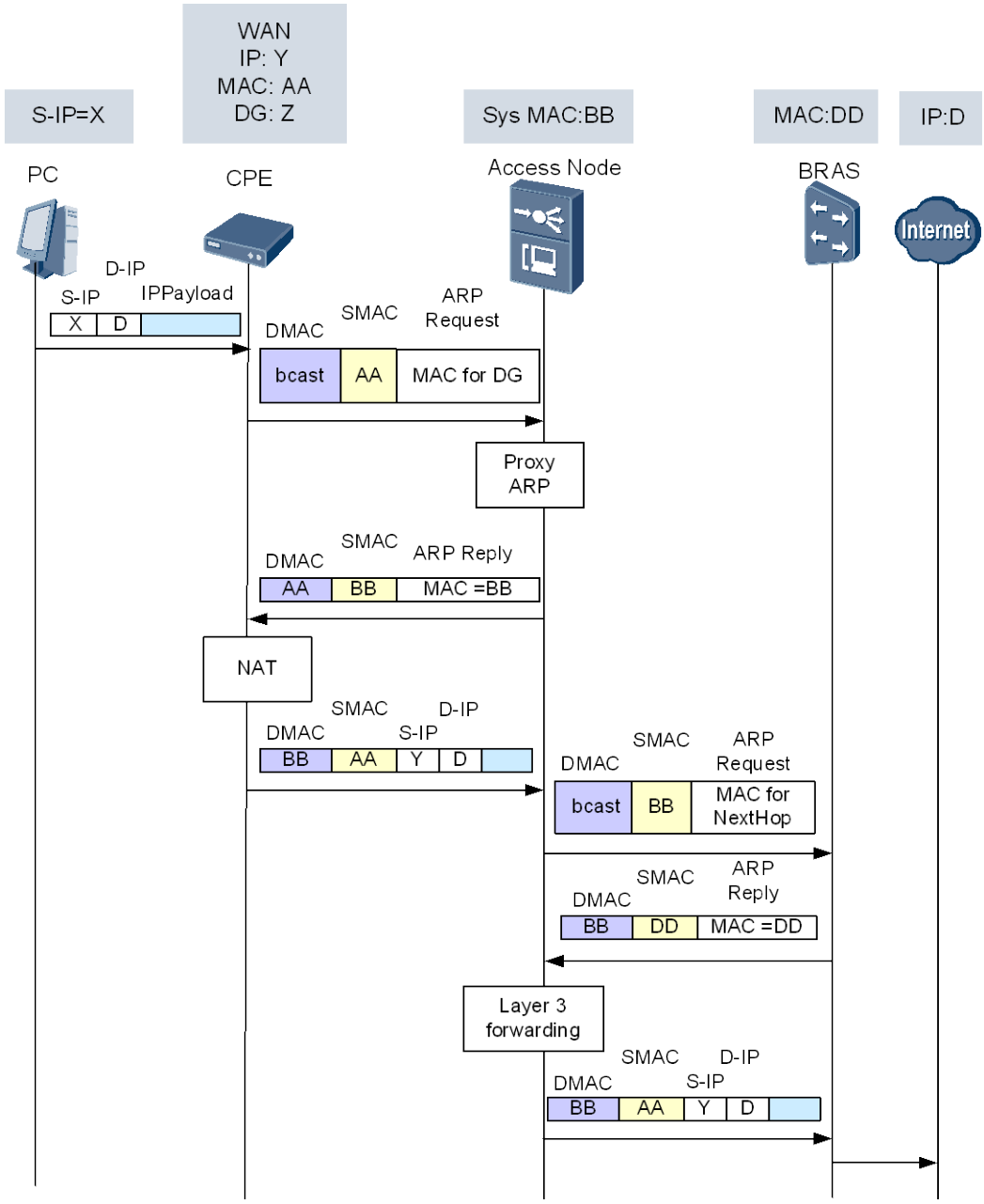
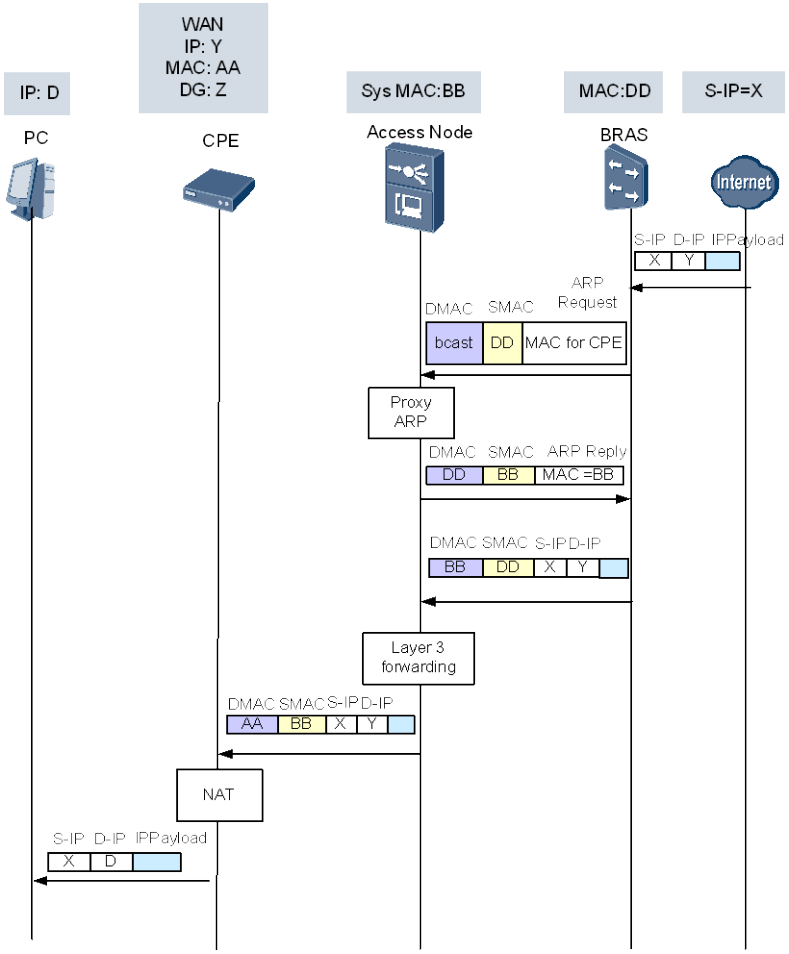


Figure 15-16 Flow of Layer 3 forwarding in the downstream direction



Key Points of the Feature

VLAN-based IP-aware bridge

- VLAN-based IP-aware bridge is similar to Layer 3 forwarding but does not occupy IP addresses.
- The access node has the interface MAC address (system MAC address).
- The access node supports static routes but does not support dynamic routing protocols.
- The VLAN can be associated with the VRF (VPN instance). The routing entry and IP address take effect within a VRF.

DHCP snooping

- The access node performs DHCP Layer 2 relay to monitor the IP address application process of users and record the IP address information about users.
- User-side ARP entries are generated according to the DHCP snooping results.

Sending the source IP address of ARP requests to the network side

There are two modes.

- The first mode: sending ARP requests by using a user IP address (default mode)

NOTE

ARP requests are not sent to the next hop when a valid user IP address does not exist.

After a user goes offline (the IP address is released), the user IP address will not be used. Instead, another valid user IP address will be used.

- The second mode: sending ARP requests by using a virtual IP address or all-zero IP address (optional mode)
  - When the user RG and the access node next hop do not belong to the same subnet, some network equipment does not respond to ARP requests. In this case, ARP requests need to be sent using a virtual IP address or all-zero IP address as the source IP address.
  - Each VLAN enabled with IP-aware bridge can be configured with eight virtual IP addresses (corresponding to eight subnets).
  - When a corresponding virtual IP address is not available, 0.0.0.0 is used as the source IP address (this method is also called dummy ARP).

Proxy response to user-side and network-side ARP requests

- For user-side ARP requests (destination IP address is the user gateway, that is, the network-side equipment of the access node, such as the BRAS)

  The access node terminates user-side ARP requests and responds by using its own MAC address (system MAC address).
- For network-side ARP requests (destination IP address is the user IP address)

  The access node terminates network-side ARP requests and responds by using its own MAC address (system MAC address).

User-side ARP interoperation

- By default, the users in the same VLAN do not interoperate with each other.
- After global ARP proxy is enabled, users can interoperate at Layer 3.

15.8.3 Configuring the IP-aware Bridge

After the IP-aware bridge function is enabled on the MA5600T/MA5603T/MA5608T, in the upstream direction of the MA5600T/MA5603T/MA5608T, data can be forwarded to different upper-layer devices according to the destination IP address and the configured static route. With the IP-aware bridge function enabled, the MA5600T/MA5603T/MA5608T features the ARP proxy function: shielding the MAC address of the network-side device for the user side and shielding the user-side MAC address for the network side.

Context

- The IP address of the VLAN L3 interface need not be configured if the IP-aware bridge function is enabled for the VLAN. Therefore, only the IP address of the convergence layer interface needs to be planned. The data, however, can still be forwarded at L3 on the MA5600T/MA5603T/MA5608T. This solves the problem of insufficient IP addresses.
- After the IP-aware bridge function is enabled, the system automatically performs ARP processing. The ARP proxy function between users, however, is not supported. To enable the communication between isolated ports in the same broadcast domain or ports in different broadcast domains, run the **arp proxy** command to enable the ARP proxy function of the user side.
- The VLAN L3 interface has no IP address after the IP-aware bridge function is enabled for the VLAN. Therefore, the dynamic routing protocol cannot be used and only static

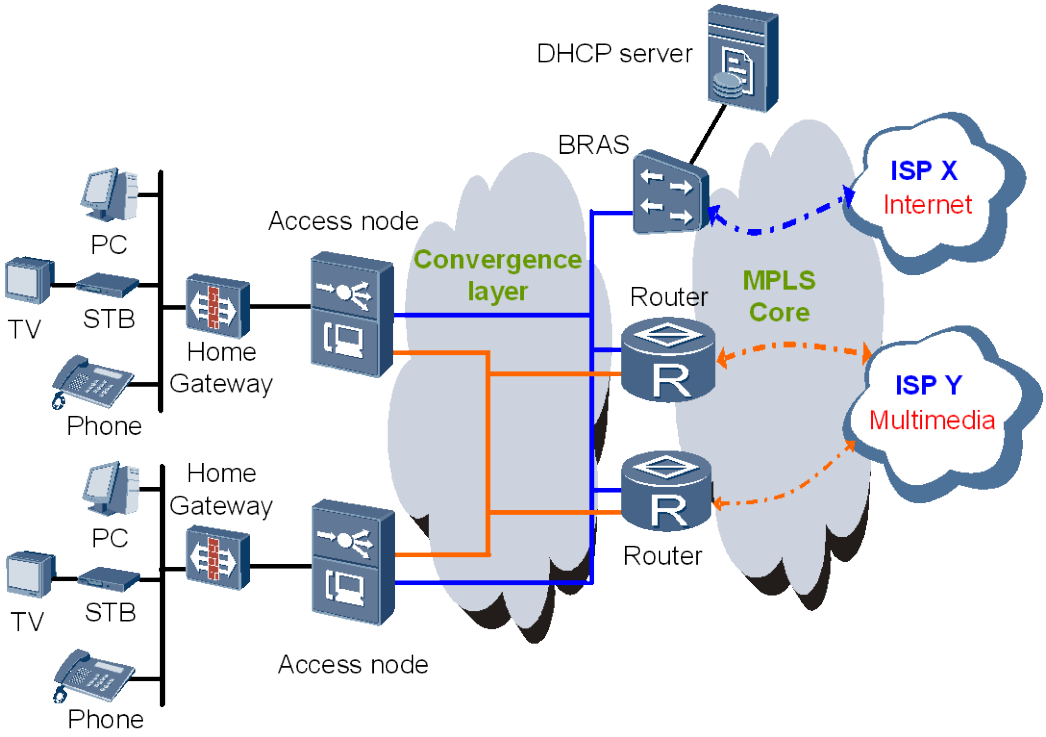
routes can be configured. If the static route of the configured IP-aware bridge feature conflicts with the static route of a normal L3 interface, the first configured static route takes effect. It is recommended that you ensure no conflict between configurations and delete the original configuration if necessary.

Networking

Figure 15-17 shows an example network of the IP-aware bridge function.

Internet access service are in ISP X and Unicast services such as VOD are in ISP Y. To differentiate the Internet access service and VOD service, private routes need to be planned. The data of these two services are forwarded through different routes.

Figure 15-17 Example network of the IP-aware bridge function



Procedure

(Optional) Create a VPN instance.

When it need to allocate different VRF (Virtual Route Forward) for different ISP, run the command **ip vpn-instance** to create a VPN instance or into the VPN instance mode. You can configure related parameters in this mode. If not configured, all the IP forwarding will be in the default VRF.

Step 1 Create a VLAN and add the upstream port to the VLAN.

Run the **vlan** command to create a VLAN, and then run the **port vlan** command to add an upstream port to the VLAN. The VLAN can be a smart VLAN, MUX VLAN, or standard VLAN. The attribute of the VLAN must be common.

Step 2 Enable the IP-aware bridge function.

Run the **ip-aware vlan** command to enable the IP-aware bridge function of the VLAN. The VLAN here corresponds to the SVLAN of the service port.

NOTE
In the VLAN whose IP-aware bridge function is enabled, the VLAN L3 interface cannot be configured; in the VLAN whose L3 interface is created, the IP-aware bridge cannot be enabled.

Step 3 Configure the obtaining mode of the source IP address in the ARP request.

In the VLAN whose IP-aware bridge function is enabled, the MAC address of the gateway needs to be obtained from the ARP request.

The obtaining mode of the source IP address in the ARP request can be configured through the **ip-aware vlan** *vlanid* **source-ip-mode** { **client-ip** | **virtual-ip** } command.

- **client-ip**: Configures the obtaining mode of the source IP address as client IP address. When the ARP request is sent, the user IP address is used as the source IP address of the ARP request. In certain networks, the upper-layer device responds only when the source IP address in the sent ARP request is the user IP address. In this case, select this mode according to the requirement of the upper-layer device. This is the default mode of the system.
- **virtual-ip**: Configures the obtaining mode of the source IP address as virtual IP address. The server for the Internet access service and that for the VOD service may be in different network segments, but the sent ARP request must reach gateways of the two servers. Therefore, the virtual IP address and the gateway address must be configured in the same network segment so that the ARP request can be transmitted correctly. You can run the **ip-aware vlan virtual-ip** command to configure the virtual IP address of the VLAN as the source IP address of the ARP request.

NOTE
Each VLAN whose IP-aware bridge function is enabled can be configured with multiple virtual IP addresses (because multiple next hops may be in different network segments). Each VLAN supports up to eight virtual IP addresses.

Step 4 (Optional)configure the period for sending the VLAN ARP request.

You can also run the **ip-aware vlan arp-send-period** command to configure the period for sending the VLAN ARP request. By default, the ARP request is sent every 180s. You can adjust the period according to actual requirements.

Step 5 Configure the IP-aware bridge static route.

Because the VLAN L3 interface has no IP address, the dynamic routing protocol cannot be used and only static routes can be configured. Configure routing entries for access nodes so that packets can be forwarded to identified IP address according to the destination IP address and routing information of the packet.

Run the **ip-aware route-static** command to configure the IP-aware bridge function. You can configure a default route for the Internet access service (next hop 0.0.0.0) and another private route for the VOD service.

Step 6 Create a service port.

Run the **service-port** command to create a service port to establish the service channel between the user and the MA5600T/MA5603T/MA5608T.

----End

Example

Assumption:

- Gateway address to ISP X (Internet access service): 10.1.1.2.
- Gateway address to ISP Y (multimedia service): 10.1.1.250, server network segment: 192.168.1.0/24.

Configure the IP-aware bridge function for VLAN 100 with the default VRF, configure the upstream port to 0/19/0, and configure the obtaining mode of the source IP address in the ARP request as client IP address. The user adopts the GPON access mode through port 0/18/0, adopts the default value for the period of sending the ARP request (no need to configure), and creates a private route to ISP X and ISP Y. To perform these operations, do as follows:

```
huawei(config)#vlan 100 smart
huawei(config)#port vlan 100 0/19 0
huawei(config)#ip-aware vlan 100
huawei(config)#ip-aware vlan 100 source-ip-mode client-ip
huawei(config)#ip-aware route-static 0.0.0.0 0 vlan 100 10.1.1.2
huawei(config)#ip-aware route-static 192.168.1.0 24 vlan 100 10.1.1.250
huawei(config)#service-port 100 vlan 100 gpon 0/18/0 gemport 128 multi-service
user-vlan 100 rx-cttr 6 tx-cttr 6
```

15.8.4 IP-aware Bridge Reference Standards and Protocols

None

16 Routing

About This Chapter

Routing is a common term used for describing the path through which the packets from a host in a network travel to a host in another network.

16.1 Introduction to Routing

Definition

Routing is a common term used for describing the path through which the packets from a host in a network travel to a host in another network.

Routers send packets on the Internet. A router selects a suitable path in a network according to the destination address included in a received packet, and sends the packet to the next router on the path. In this way, the packet travels over the Internet Until it reaches the destination host.

Purpose

The access equipment, serving as a basic element in the entire telecom network, must support the functions of remote operation, management and maintenance on the equipment itself.

With the development of small-size access equipment that can be managed remotely, the access equipment needs to feature the functions of a BRAS, such as allocation of network addresses and user management. In this way, the access equipment must support the routing feature.

A MA5600T/MA5603T/MA5608T can also serve as a router.

16.2 Routers

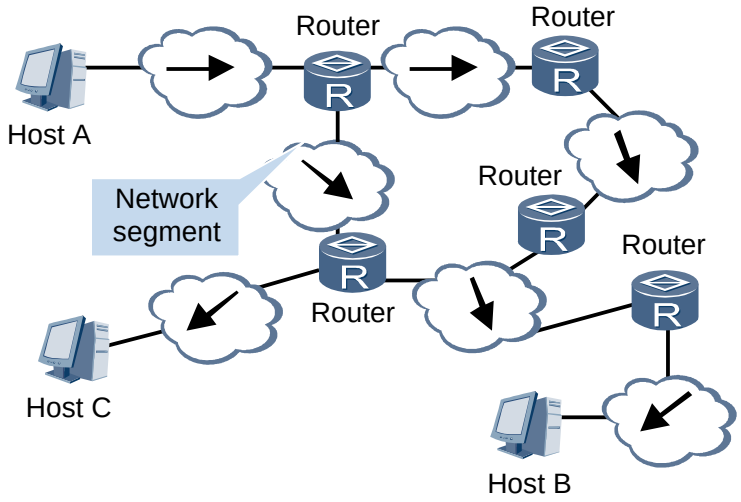
In the Internet, network connecting devices control traffic and ensure the quality of data transmission on the network. Common network connecting devices include hubs, bridges, switches and, routers.

As a typical network connection device, a router is used to select routes and forward packets. According to the destination address in the received packet, a router selects a proper path, which has single-hop or multiple hops in it, to send the packet to the next router. The last router is responsible for sending the packet to the destination host. In addition, the router can select an optimal path to transmit data.

The router logically takes the path through which a packet travels from the network ingress to the network egress as a route unit; this is called a hop. The path that a hop covers is called a network segment. The hop count from a router to its directly connected network is zero, and to a network through another router, is one. The remaining number of hops required for the route can be deduced by analogy. If a router is connected to another router through a network, that is, a network segment exists between the two routers, the two routers are considered as adjacent routers on the Internet. This connection between routers is independent of the physical links that constitute each network segment.

As shown in Figure 16-1, the packets from Host A travel through three networks and two routers until they reach Host C and the hop count is three.

Figure 16-1 Working principle of routers



Virtual route forwarding instance (VRF) is a mechanism in which a device works as multiple virtual routing devices. The MA5600T/MA5603T/MA5608T categorizes VRFs by VLANs to provide L3VPN solutions. All the packets or related protocols on the Layer 3 interface of a VRF are processed only in this VRF, which is unrelated to other VRFs. In this way, the services or users can be isolated, and the IP addresses can be saved.

16.3 Routing Table and FIB Table

Each router maintains one routing table and one FIB table at least. A router uses the routing table to select routes, and uses the FIB table to guide packet forwarding. In Equal and Weighted Cost Multi-Path (ECMP), according to different states of the network, the traffic to the same destination network can be distributed among multiple equal-cost paths to reduce the network load or to implement the link backup function.

- Routes discovered by the various routing protocols are stored in the routing table. The routes in the routing table are divided, according to their sources, into the following types:
  - Directly connected route or interface route: is the route discovered by the link layer protocols.
  - Static route: is the route manually configured by the network administrator.
  - Dynamic route: is the route discovered by dynamic routing protocols.
- Each entry in the FIB table contains the physical or logical interface through which a packet is sent to a network segment or host to reach the next router. An entry also indicates whether the packet can be sent directly to a destination host in a directly connected network.

Routing Table

The routing table is key for forwarding packets. The route entries in the table are used for the following:

- Through which physical interface of the router a packet can be forwarded to a specific subnet or host so as to reach the next router along the path.
- Whether the packet can be sent to the destination host in an interconnected network without passing through other routers.

Each router maintains the protocol routing table for each type of protocol and a local core routing table (or routing management table).

- Protocol routing table

  A protocol routing table stores the routing information discovered by the protocol. A routing protocol can import and advertise the routes that are discovered by other protocols. For example, if a router that runs the Open Shortest Path First (OSPF) protocol needs to use OSPF to advertise direct routes, static routes, or Intermediate System-Intermediate System (IS-IS) routes, the router must import the routes into the OSPF routing table.
- Local core routing table

  A router uses the local core routing table to store protocol routes and preferred routes. The router then sends the preferred routes to the FIB table to guide packet forwarding. The router selects routes according to the priorities of protocols and costs stored in the routing table. To view the local core routing table of a router, run the **display ip routing-table** command.

  A router maintains a local core routing table for each Virtual route forwarding instance (VRF).

The key entries of the routing table are shown in Table 16-1.

Table 16-1 Key entries of the routing table

Entry	Description
Destination	The destination address is a 32-bit character that labels the destination IP address or destination network of an IP packet.
Mask	The mask is used with the destination address to identify the subnet address of the destination host or router. The network mask is composed of several consecutive 1s. These 1s can be

Entry	Description
	expressed in either the dotted decimal notation or the number of consecutive 1s in the mask. For example, the network mask can be expressed either as 255.255.255.0 or 24. The network address of the destination host or router is obtained through the "AND" operation on the destination address and network mask. For example, if the destination address is 1.1.1.1 and the mask is 255.255.255.0, the address of the network where the host or router resides is 1.1.1.0.
Proto	Indicates the protocol through which routes are learned.
Pre	Indicates the preference added to the IP routing table for a route. To the same destination, multiple routes with different next hops and outgoing interfaces exist. The routes in the table are those discovered by different routing protocols or tare the manually configured static routes. The router selects the route with the highest preference (the smallest value) as the optimal route. For more information on the preference of each protocol, see Table 16-4 in the *Route Protocols*.
Cost	Indicates the route cost. When multiple routes to the same destination have the same preference, the route with the lowest cost is selected as the optimal route. **NOTE** The Preference value is used to compare the preferences of various routing protocols, while the Cost value is used to compare the preferences of different routes of the same routing protocol.
NextHop	Indicates the IP address of the next device that an IP packet passes through.
Interface	Indicates the outgoing interface through which an IP packet is forwarded.

Based on the destination, routes can be classified as:

- Subnet route: Its destination is a subnet.
- Host route: Its destination is a host.

Based on the connection between the destination and the router, routes can be classified as:

- Direct route: Its destination network is directly connected to the router.
- Indirect route: Its destination network is not directly connected to the router.

To avoid large routing tables, a default route can be assigned. Once a packet fails to find a dedicated route in the routing table, the default route is selected for forwarding the packet.

Figure 16-2 and Table 16-2 shows some interconnected networks. The digits in each network represent the IP address of the network. Router 8 is connected to three networks. Therefore, it has three IP addresses and three physical ports.

Figure 16-2 Interconnected networks

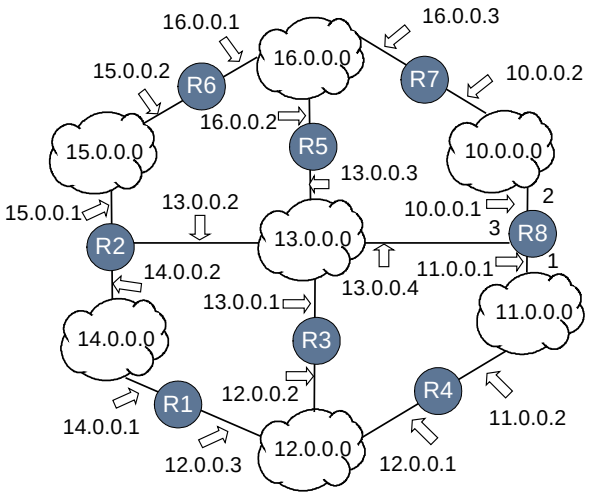


Table 16-2 Routing table of Router 8

Subnet of the Destination Host	Forward or Forward from...	Through Port...
10.0.0.0	Directly	2
11.0.0.0	Directly	1
12.0.0.0	11.0.0.2	1
13.0.0.0	Directly	3
14.0.0.0	13.0.0.2	3
15.0.0.0	10.0.0.2	2
16.0.0.0	10.0.0.2	2

FIB Table

After route selection, routers send the active routes in the routing table to the FIB table. When a router receives a packet, the router searches the FIB table for the optimal route to forward the packet.

The router performs the "AND" operation on the destination address in the packet and the network mask of each entry in the FIB table. The router then compares the result of the "AND" operation with the entries in the FIB table to find a match. The router chooses the optimal route to forward packets according to the best or "longest" match.

As an example, a certain router has the following brief routing table:

```
Routing Tables:
Destination/Mask    Proto   Pre  Cost      Flags NextHop         Interface
      0.0.0.0/0    Static   60   0           D    120.0.0.2       Vlanif10
      8.0.0.0/8    RIP     100  3           D    120.0.0.2       Vlanif10
```

```
     9.0.0.0/8   OSPF    10   50     D    20.0.0.2       Vlanif10
     9.1.0.0/16  RIP     100  4      D    120.0.0.2      Vlanif20
    20.0.0.0/8   Direct  0    0      D    20.0.0.1       Vlanif20
```

After receiving a packet that carries the destination address 9.1.2.1, the router searches the following table:

```
 FIB Table:
 Total number of Routes : 5
Destination/Mask   Nexthop        Flag TimeStamp    Interface        TunnelID
0.0.0.0/0         120.0.0.2      SU   t[37]        Vlanif10               0x0
8.0.0.0/8         120.0.0.2      DU   t[37]        Vlanif10               0x0
9.0.0.0/8         20.0.0.2       DU   t[9992]      Vlanif10               0x0
9.1.0.0/16        120.0.0.2      DU   t[9992]      Vlanif20               0x0
20.0.0.0/8        20.0.0.1       U    t[9992]      Vlanif20               0x0
```

Then the router performs the "AND" operation on the destination address 9.1.2.1 and the masks 0, 8, 16 to obtain the network segment addresses: 0.0.0.0/0, 9.0.0.0/8, and 9.1.0.0/16. The three addresses match three entries in the table. The router chooses the 9.1.0.0/16 entry because it is the longest match. The router then forwards the packet through Pos 2/0/0 for the 9.1.0.0/16 entry.

16.4 Routing Protocols

The MA5600T/MA5603T/MA5608T supports the configuration of static routes and the dynamic routing protocols such as RIP, OSPF, IS-IS, and BGP. The MA5600T/MA5603T/MA5608T manages the static and dynamic routes in a unified manner. The static routes and the routes discovered by the routing protocols can be shared.

- Static routes can be easily configured on a system and have lower system requirements. Static routes are applicable to simple, stable, and small-scale networks. Static routes, however, cannot automatically adapt to changes in the network topology, so they must be manually configured.
- On the other hand, dynamic routing protocols use routing algorithms to automatically adapt to changes in network topology. Dynamic routes are applicable to the network that is equipped with Layer 3 devices. The dynamic route configuration, however, has a higher requirement (such as large memory capacity) for system performance and occupies more network resources.

Classification of Dynamic Routing Protocols

Dynamic routing protocols are classified according to the following factors:

Routing protocols are classified according to the application range:

- Interior Gateway Protocol (IGP): runs inside an AS, such as RIP, OSPF, and IS-IS.

Table 16-3 Differences among the three typical IGPs

Item	RIP	OSPF	IS-IS
Protocol type	IP layer protocol	IP layer protocol	Link layer protocol

Item	RIP	OSPF	IS-IS
Application scope	Applies to small networks with simple architectures, such as campus networks.	Applies to medium-sized networks, such as enterprise networks.	Applies to large networks, such as Internet service provider (ISP) networks.
Routing algorithm	Uses a distance-vector algorithm and exchanges routing information over the User Datagram Protocol (UDP).	Uses the shortest path first (SPF) algorithm to generate a shortest path tree (SPT) based on the network topology, calculates shortest paths to all destinations, and exchanges routing information over IP.	Uses the SPF algorithm to generate an SPT based on the network topology, calculates shortest paths to all destinations, and exchanges routing information over IP. The SPF algorithm runs separately in Level-1 and Level-2 databases.
Route convergence speed	Slow	Less than 1 second	Less than 1 second
Scalability	Not supported	Supported by partitioning a network into areas	Supported by defining levels

- Exterior Gateway Protocol (EGP): runs between different ASs, such as BGP.

Routing protocols are classified according to the type of algorithm they use:

- Distance-Vector Routing Protocol: includes RIP and BGP (BGP is also called Path-Vector).
- Link-State Routing Protocol: includes OSPF and IS-IS.

The current route to a specific destination at a specific moment can only be determined by one routing protocol. Each routing protocol (including the static routing protocol) is allocated a preference. When multiple route sources exist, the route discovered by the routing protocol with the highest preference becomes the current route.

The smaller the value, the higher the preference. In this table, "0" indicates the direct route, and "255" indicates any route from an untrusted source.

You can define the priorities for all dynamic routing protocols except the direct route (DIRECT) and the BGP (IBGP, EBGP). In addition, the priorities of any two static routes can be different.

Routing Protocols and Routing Preference

Routing protocols (including the static route) can learn different routes to the same destination, but not all routes are optimal. Only one routing protocol at one time determines the optimal route to a destination. To select the optimal route, each routing protocols (including the static route) is configured with a preference (the smaller the value, the higher the preference). When multiple routing information sources coexist, the route with the highest preference is selected as the optimal route (the smaller the value is, the higher the preference is).

Table 16-4 lists various routing protocols and the default priorities of the routes discovered by them. "0" indicates the direct route, and "255" indicates any route from an untrusted source.

Table 16-4 Default priorities of routing protocols

Routing Protocol or Route Type	Default Routing Preference
DIRECT	0
OSPF	10
IS-IS	15
INTERNAL EIGRP	50
STATIC	60
RIP	100
OSPF ASE (AS-External)	150
OSPF NSSA (Not-So-Stubby Area)	150
EXTERNAL EIGRP	160
IBGP	255
EBGP	255
UNKNOWN	255

You can define the priorities for all dynamic routing protocols except the direct route (DIRECT) and the BGP (IBGP, EBGP). In addition, the priorities of any two static routes can be different.

If different routing protocols are configured with the same preference, the system determines which routes discovered by these routing protocols become the preferred routes through an internal preference. Table 16-5 shows the internal preferences of routing protocols.

Table 16-5 Internal preferences of routing protocols

Routing Protocol or Route Type	Internal Routing Preference
DIRECT	0
OSPF	10
IS-IS Level-1	15
IS-IS Level-2	18
STATIC	60
RIP	100
OSPF ASE (AS-External)	150
OSPF NSSA (Not-So-Stubby Area)	150

Routing Protocol or Route Type	Internal Routing Preference
IBGP	200
EBGP	20

For example, two routes, an OSPF route and a static route, can reach the destination 10.1.1.0/24, and the preferences of both routes are set to 5. In this case, the VRP determines the optimal route according to the internal preferences listed in Table 16-5. The internal preference value 10 of OSPF is higher than the internal preference value 60 of the static route. Therefore, the system selects the route discovered by OSPF as the optimal route.

Route Sharing through Route Policy

Different routing protocols can find different routes as they use different algorithms. A routing protocol might need to import routes discovered by other protocols to diversify its own routes. The MA5600T/MA5603T/MA5608T supports importing the routes discovered by one protocol to another protocol. Each protocol has its own route importing mechanism.

However, a protocol only needs to import qualified routes by setting attributes of the routes to be imported. To realize a route policy, you must define the attributes of the routes to which the route policy is to be applied, such as the destination address, and the address of the router distributing routes. You can define the matching rules in advance so that they can be applied in a route policy for route distribution, reception and importing.

The two applications of the routing policy are as follows:

- When importing routes discovered by other protocols, a routing protocol can apply this filter to obtain the required routes.
- When transmitting or receiving routes, a routing protocol can apply the filter so that only the required routes are transmitted or received.

16.5 Static Routes

16.5.1 Introduction to Static Routes

Definition

Static routes need to be manually configured by the administrator.

Purpose

On a simple network, the administrator just needs to configure static routes so that the network can run properly. Properly configuring and using static routes can improve network performance and guarantee the required bandwidth for important applications.

16.5.2 Components of Static Routes

On the MA5600T/MA5603T/MA5608T, you can run the **ip route-static** command to configure a static route, which consists of the following:

- Destination Address and Mask
- Outbound Interface and Next-Hop Address

Destination Address and Mask

In the **ip route-static** command, the IPv4 address is expressed in dotted decimal notation. The mask is expressed in dotted decimal notation or represented by the mask length (the number of consecutive 1s in the mask).

Outbound Interface and Next-Hop Address

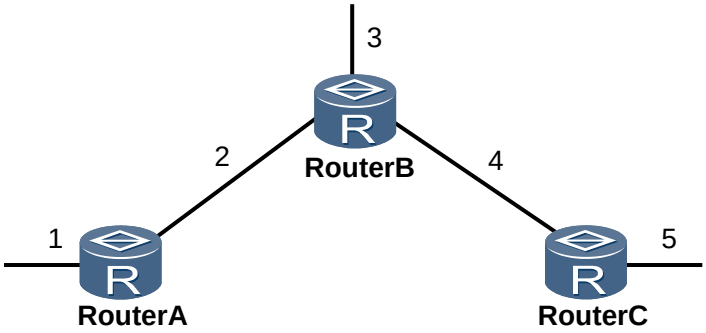
When you configure a static route, you can specify an outbound interface, a next-hop address, or both as required.

Actually, each routing entry requires a next-hop address. Before sending a packet, a device needs to use the longest match rule to search its routing table for the route that matches the destination address in the packet. The device can find the associated link layer address only after the next-hop address of the packet is specified.

16.5.3 Applications of Static Routes

As shown in Figure 16-3, the network topology of static routes is simple, and network communication can be implemented through static routes. In this application, you must specify an address for each physical network, identify indirectly connected physical networks for each Router, and configure static routes for the indirectly connected physical networks.

Figure 16-3 Static routes networking



In Figure 16-3, static routes to networks 3, 4, and 5 need to be configured on Router A; static routes to networks 1 and 5 need to be configured on Router B; and static routes to networks 1, 2, and 3 need to be configured on Router C.

Default Static Route

When you run the **ip route-static** command to configure a static route, if the destination address and the mask are both set to all 0s (0.0.0.0 0.0.0.0), a default route is configured. This condition simplifies the network configuration.

In Figure 16-3, because the next hop of the packets sent by Router A to networks 3, 4, and 5 is Router B, a default route can be configured on Router A to replace the three static routes destined for networks 3, 4, and 5 in the preceding example. Similarly, only a default route

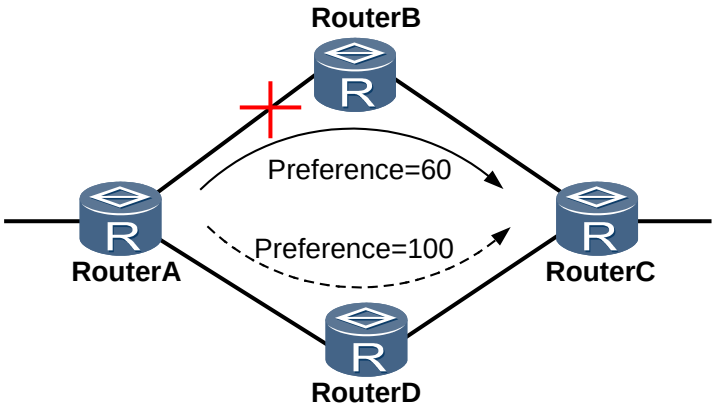
from Router C to Router B needs to be configured to replace the three static routes destined for networks 1, 2, and 3 in the example.

Floating Static Routes

Static routes can be configured with different preferences so that routing management policies can be flexibly applied. Different preferences specified for multiple routes to the same destination can be used to implement route backup. Figure 16-4 shows the networking for floating static routes.

As shown in Figure 16-4, there are two static routes from Router A to Router C. Normally, in the routing table, only the static route with the next hop being Router B is in the Active state because this route has a higher preference. The other static route with the next hop being Router D functions as a backup route. The backup route is activated to forward data only when the primary link becomes faulty. After the primary link recovers, the static route with the next hop being Router B becomes active to forward data. Therefore, the backup route is also called a floating static route. The floating static route becomes ineffective when a fault occurs on the link between Router B and Router C.

Figure 16-4 Floating static routes

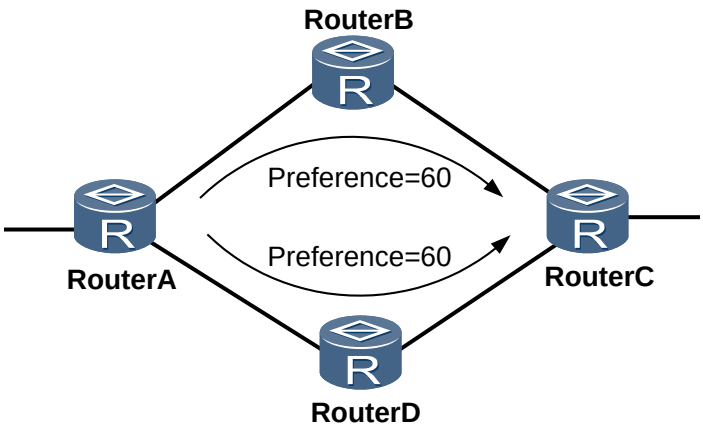


Load Balancing Among Static Routes

Specifying the same preference for multiple routes to the same destination can implement load balancing.

As shown in Figure 16-5, there are two static routes with the same preference from Router A to Router C. The two routes exist in the routing table and forward data at the same time.

Figure 16-5 Load balancing among static routes



16.5.4 Functions of Static Routes

IPv4 Static Routes

The MA5600T/MA5603T/MA5608T supports common static routes and the static routes associated with VPN instances. The static routes associated with VPN instances are used to manage VPN routes.

Attributes and Functions of IPv6 Static Routes

Similar to IPv4 static routes, IPv6 static routes need to be manually configured by the administrator. IPv6 static routes are applicable to simple IPv6 networks.

The major difference between IPv6 static routes and IPv4 static routes lies in their destination addresses and next-hop addresses. IPv6 static routes use IPv6 addresses, whereas IPv4 static routes use IPv4 addresses.

During the configuration of an IPv6 static route, if the specified destination address is ::/0 (the mask length is 0), it indicates that a default IPv6 route is configured. If the destination address of a packet fails to match any entry in the routing table, a Router selects the default IPv6 route to forward the IPv6 packet.

16.5.5 BFD for Static Routes

Unlike dynamic routing protocols, static routes do not have a detection mechanism. When a fault occurs on the network, the administrator needs to handle it. To assist the administrator, Bidirectional Forwarding Detection (BFD) for static routes can be introduced into the network to bind a static route to a BFD session. Then the BFD session can detect the status of the link where the static route resides.

After BFD for static route is configured, each static route can be bound to a BFD session.

- If the BFD session on the link of a static route detects that the link changes from Up to Down, BFD reports it to the system. Then, the system deletes the route from the IP routing table.

- When a BFD session is established on the link of a static route or the BFD session changes from Down to Up, BFD reports it to the system. Then, the system adds the route to the IP routing table.

BFD for static routes has one mode:

- Single-hop detection

  For a non-iterated static route, the configured outbound interface and next-hop address provide the information about the directly connected next hop. In this case, the outbound interface bound to the BFD session is the outbound interface of the static route, and the peer address is the next-hop address of the static route.

NOTE

If the next hop of a route is not directly reachable, the route cannot be used for packet forwarding. Based on information about the current next hop of this route, the system will calculate an actual outbound interface and an actual next hop. This process is called route iteration. In the **display ip routing-table** command output, if the **Flags** value of a route is displayed R, the route is an iterated route. Otherwise, the route is not an iterated route.

NOTE

Only IPv4 supports BFD for static routing.

16.5.6 Permanent Advertisement of Static Routes

Link connectivity determines the stability and availability of a network. Therefore, link detection plays an important role in network maintenance. BFD, as a link detection mechanism, is inapplicable to certain scenarios. For example, a simpler and more natural method is required for link detection between different Internet Service Providers (ISPs).

Permanent advertisement of static routes provides a low-cost and simple link detection mechanism and improves compatibility between Huawei devices and non-Huawei devices. If service traffic needs to be forwarded along a specified path, you can ping the destination addresses of static routes to detect the link connectivity.

After permanent advertisement of static routes is configured, the static routes that cannot be advertised are still preferred and are added to the routing table in the following cases:

- If an outbound interface configured with an IP address is specified for a static route, the static route is always preferred and added to the routing table regardless of whether the outbound interface is Up or Down.
- If no outbound interface is specified for a static route, the static route is always preferred and added to the routing table regardless of whether the static route can be iterated to an outbound interface.

In this way, you can enable IP packets to be always forwarded through this static route. The permanent advertisement mechanism provides a way for you to monitor services and detect link connectivity.

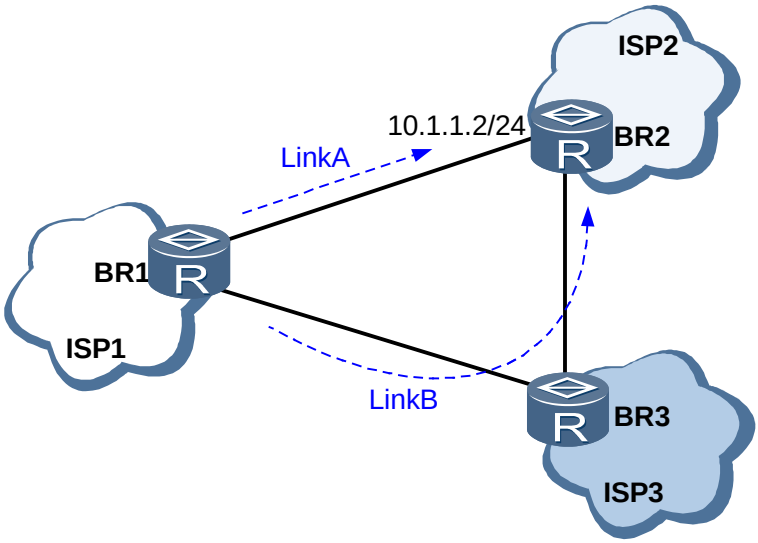
NOTE

A device enabled with this feature always stores static routes in its IP routing table, regardless of whether the static routes are reachable. If a path is unreachable, the corresponding static route may become a blackhole route.

Applications

In Figure 16-6, BR1, BR2, and BR3 belong to ISP1, ISP2, and ISP3 respectively. Between BR1 and BR2 are two links, Link A and Link B. ISP1, however, requires that service traffic be forwarded to ISP2 over Link A without traveling through ISP3.

Figure 16-6 Networking for applying permanent advertisement of static routes



The External Border Gateway Protocol (EBGP) peer relationship is established between BR1 and BR2. For service monitoring, a static route destined for the BGP peer (BR2) at 10.1.1.2/24 is configured on BR1, and permanent advertisement of static routes is enabled. The interface that connects BR1 to BR2 is specified as the outbound interface of the static route. Then, the network monitoring system periodically pings 10.1.1.2 to determine the status of Link A.

If Link A works properly, ping packets are forwarded over Link A. If Link A becomes faulty, although service traffic can reach BR2 over Link B, the static route is still preferred because its preference is higher. Therefore, ping packets are still forwarded over Link A, but packet forwarding fails. This scenario is also applicable to BGP packets. That is, a link fault causes the BGP peer relationship to be interrupted. The monitoring system detects service faults as returned in the ping result and prompts maintenance engineers to rectify the faults before services are affected.

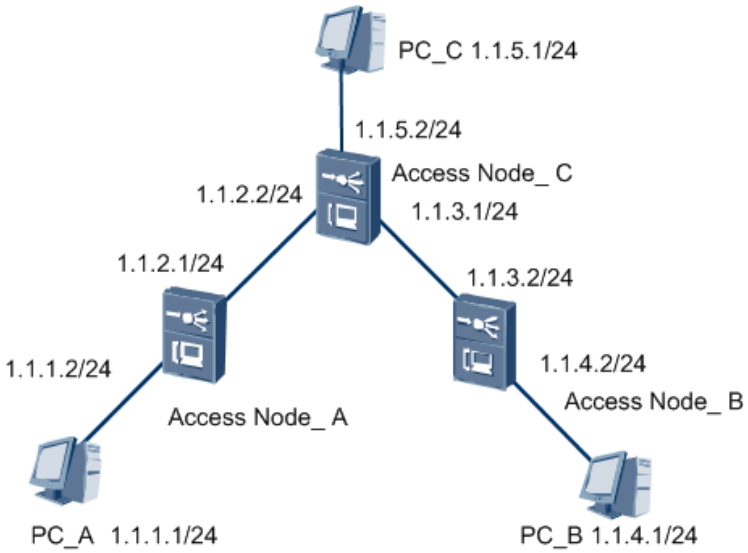
16.5.7 Configuration Example of the IPv4 Static Route

This topic describes how to manually add the IPv4 static route to implement the interconnection between MA5600T/MA5603T/MA5608T.

Service Requirements

In this example network, MA5600T/MA5603T/MA5608T_A, MA5600T/MA5603T/MA5608T_B, and MA5600T/MA5603T/MA5608T_C have the routing function. It is expected that after the configuration, any two PCs can communicate with each other.

Figure 16-7 Example network for configuring the IPv4 static route



Procedure

Configure the IP address of the Layer 3 interface.

The configurations for the three MA5600T/MA5603T/MA5608T devices are the same. The configuration of the MA5600T/MA5603T/MA5608T is considered as an example.

```
huawei(config)#vlan 2 smart
huawei(config)#port vlan 2 0/19 0
huawei(config)#interface vlanif 2
huawei(config-if-vlanif2)#ip address 1.1.1.2 24
huawei(config-if-vlanif2)#ip address 1.1.2.1 24 sub
huawei(config-if-vlanif2)#quit
```

Step 1 Configure IPv4 static routes.

1. Configure an IPv4 static route for MA5600T/MA5603T/MA5608T_A.

```
huawei(config)#ip route-static 1.1.5.0 255.255.255.0 1.1.2.2
huawei(config)#ip route-static 1.1.4.0 255.255.255.0 1.1.2.2
```

2. Configure an IPv4 static route for MA5600T/MA5603T/MA5608T_B.

```
huawei(config)#ip route-static 1.1.5.0 255.255.255.0 1.1.3.1
huawei(config)#ip route-static 1.1.1.0 255.255.255.0 1.1.3.1
```

3. Configure IPv4 static routes for MA5600T/MA5603T/MA5608T_C.

```
huawei(config)#ip route-static 1.1.1.0 255.255.255.0 1.1.2.1
huawei(config)#ip route-static 1.1.4.0 255.255.255.0 1.1.3.2
```

Step 2 Configure the host gateways.

1. Configure the default gateway of Host A to 1.1.1.2.
2. Configure the default gateway of Host B to 1.1.4.2.
3. Configure the default gateway of Host C to 1.1.5.2.

Step 3 Save the data.

```
huawei#save
```

----End

Result

After the configuration, an interconnection can be set up between all the hosts and between all the MA5600T/MA5603T/MA5608T devices. Run the **ping** and **tracert** command to check the network connectivity.

Run the **display ip routing-table** command to query the IPv4 routing table which contains the static routing information that is configured.

Configuration File

Configuration example of MA5600T/MA5603T/MA5608T_A.

```
vlan 2 smart
port vlan 2 0/19 0
interface vlanif 2
ip address 1.1.1.2 24
ip address 1.1.2.1 24 sub
quit
ip route-static 1.1.5.0 255.255.255.0 1.1.2.2
ip route-static 1.1.4.0 255.255.255.0 1.1.2.2
```

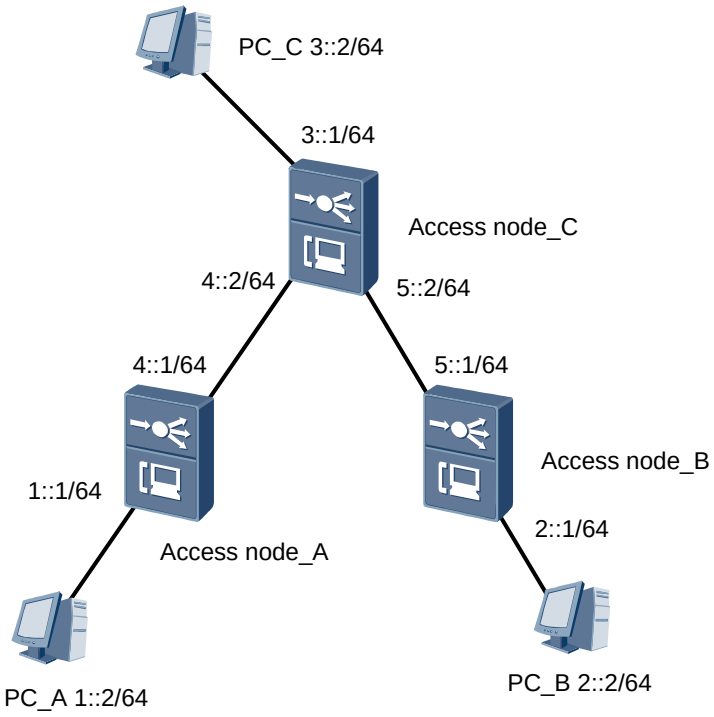
16.5.8 Configuration Example of the IPv6 Static Route

This topic describes how to manually add the IPv6 static route to implement the interconnection between access devices.

Service Requirements

In this example network, three access nodes have the routing function. It is expected that after the configuration, any two PCs can communicate with each other.

Figure 16-8 Example network for configuring the IPv6 static route



Procedure

Configure the IPv6 address of the Layer 3 interface.

The configurations for the three access devices are the same. The configuration of the Access node_A is considered as an example.

```
huawei(config)#ipv6
huawei(config)#vlan 2 smart
huawei(config)#port vlan 2 0/19 0
huawei(config)#interface vlanif 2
huawei(config-if-vlanif2)#ipv6 enable
huawei(config-if-vlanif2)#ipv6 address 1::1 64
huawei(config-if-vlanif2)#quit
huawei(config)#vlan 3 smart
huawei(config)#port vlan 3 0/19 1
huawei(config)#interface vlanif 3
huawei(config-if-vlanif3)#ipv6 enable
huawei(config-if-vlanif3)#ipv6 address 4::1 64
huawei(config-if-vlanif3)#quit
```

Step 1 Configure IPv6 static routes.

1. Configure IPv6 static route for Access node_A.

   ```
   huawei(config)#ipv6 route-static :: 0 4::2
   ```

2. Configure IPv6 static route for Access node_B.

   ```
   huawei(config)#ipv6 route-static :: 0 5::2
   ```

3. Configure IPv6 static routes for Access node_C.

```
huawei(config)#ipv6 route-static 1::1 64 4::1
huawei(config)#ipv6 route-static 2::1 64 5::1
```

Step 2 Configure the host gateways.

1. Configure the default gateway of Host A to 1::1.
2. Configure the default gateway of Host B to 2::1.
3. Configure the default gateway of Host C to 3::1.

Step 3 Save the data.

```
huawei#save
```

----End

Result

After the configuration, an interconnection can be set up between all the hosts and between all the access devices. Run the **ping ipv6** and **tracert ipv6** commands to query the network connectivity.

Run the **display ipv6 routing-table** command to query the IPv6 routing table which contains the static routing information that is configured.

Configuration File

Configuration example of Access node_A.

```
ipv6
vlan 2 smart
port vlan 2 0/19 0
interface vlanif 2
ipv6 enable
ipv6 address 1::1/64
quit
vlan 3 smart
port vlan 3 0/19 1
interface vlanif 3
ipv6 enable
ipv6 address 4::1/64
quit
ipv6 route-static :: 0 4::2
```

16.5.9 References

None.

16.6 RIP

Routing Information Protocol (RIP) is a dynamic routing protocol based on the V-D algorithm. Based on RIP, the routing information is exchanged through UDP data packets. RIP is a simple Interior Gateway Protocol (IGP). It is mainly used in the smaller network, such as the campus network or the regional network with a simple topology. RIP is not recommended for a larger network in a complex environment.

16.6.1 Introduction to RIP

Definition

As a simple Interior Gateway Protocol, Routing Information Protocol (RIP) is mainly used in small-scale and simply structured networks such as campus and regional networks. RIP is not suitable for complex environments or large-scale networks.

RIP is based on the Distance-Vector (DV) algorithm. It exchanges routing information through User Datagram Protocol (UDP) packets. RIP uses he port number 520.

RIP employs Hop Count (HC) to measure the distance to the destination. The distance is called the metric value. In RIP, the default HC from a router to its directly connected network is 0, and the HC from a router to a network that is reachable through another router is 1, and so on. That is to say, the HC equals the number of routers passed from the local network to the destination network. To speed up the convergence, RIP defines the HC as an integer that ranges from 0 to 15. An HC 16 or greater is defined as infinity, that is, the destination network or the host is unreachable. For this reason, RIP is not applied to large-scale networks.

To improve performance and to prevent routing loops, RIP supports the features split horizon, poison reverse, and triggered updates.

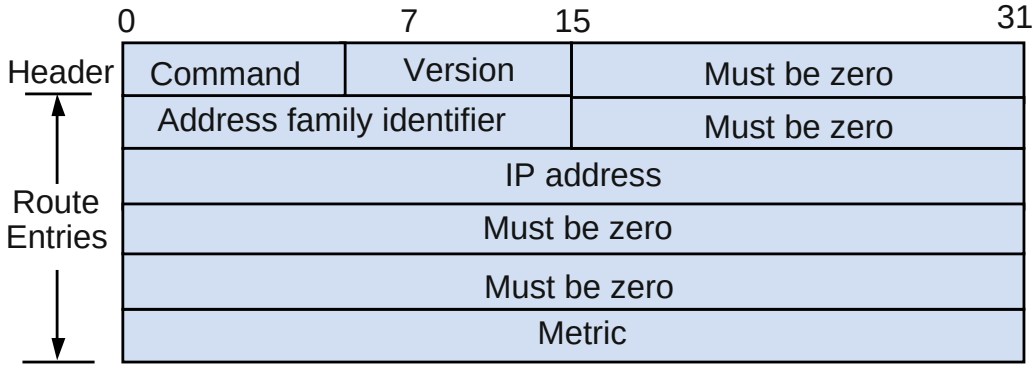
Purpose

As one of the earliest IGP protocols, RIP is widely used in small-scale networks that support RIP. The implementation of RIP is simple. The configuration and maintenance of RIP are easier than that of the Open Shortest Path First (OSPF) and Intermediate System-to-Intermediate System (IS-IS) protocols.

16.6.2 RIP-1

RIP version 1 (RIP-1) is a classful (as opposed to classless) routing protocol. It supports the advertisement of protocol packets only in broadcast mode. Figure 16-9 shows the packet format. A RIP packet can carry a maximum of 25 entries. RIP is based on UDP, and a RIP-1 data packet cannot be longer than 512 bytes. The RIP-1 protocol packet does not carry any mask, so it can identify only the routes of the natural network segment such as Class A, Class B, and Class C. Therefore, RIP-1 does not support route aggregation or discontinuous subnet.

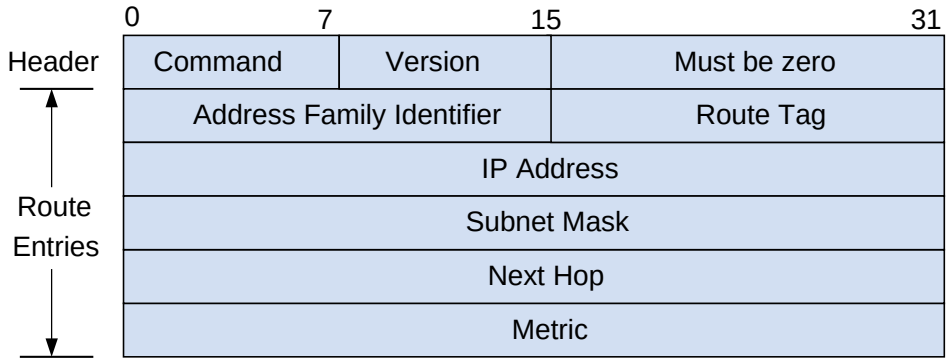
Figure 16-9 RIP-1 packet format



16.6.3 RIP-2

RIP version 2 (RIP-2), is a classless routing protocol. Figure 16-10 shows the packet format.

Figure 16-10 RIP-2 packet format



Compared with RIP-1, RIP-2 has the following advantages:

- Supports route tag and can flexibly control routes on the basis of the tag in the routing policy.
- Has packets that contain mask information and support route aggregation and Classless Inter-domain Routing (CIDR).
- Supports the next hop address and can select the optimal next hop address in the broadcast network.
- Uses multicast routes to send update packets. Only RIP-2 routers can receive protocol packets. This reduces the resource consumption.

16.6.4 Timers

RIP mainly uses the following three timers:

- Update timer: triggers the sending of update packets every 30s.
- Age timer: sets and keeps track of the 180-second time limit. If a RIP router does not receive an update packet from any of its neighbors within the aging time, the RIP router detects the route as unreachable.
- Garbage-Collect timer: determines when to delete a packet entry. If the route is no longer valid after the timer expires, the entry is removed from the RIP routing table.

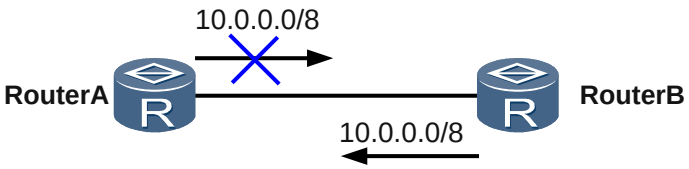
The three timers work together in the following way:

The advertisement of RIP routing update is triggered by the update timer every 30 seconds. Each entry is associated with the age timer and garbage-collect timer. When a route is learned and added in the routing table, the age timer is initialized. If no Update packet is received from the neighbor for 180 seconds, the metric value of the route is set to 16 (to specify the route as unreachable). At the same time, the garbage-collect timer is initialized. If no Update packet is received for 120 seconds, the entry is deleted after the garbage-collect timer expires.

16.6.5 Split Horizon

The principle of split horizon is that a route learned by RIP on an interface is not sent to neighbors from the interface. This reduces bandwidth consumption and avoids route loops.

Figure 16-11 Networking for split horizon

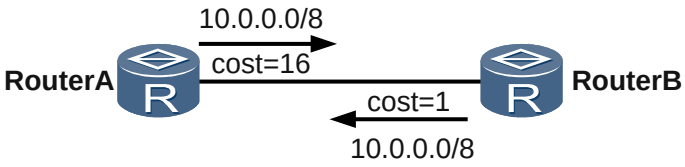


In Figure 16-11, Router B sends a route to 10.0.0.0 to Router A and Router A does not send the route back to Router B.

16.6.6 Poison Reverse

The principle of poison reverse is that RIP sets the cost of the route learned from an interface of a neighbor to 16 (to specify the route as unreachable) and then sends the route from the interface back to the neighbor. In this way, RIP can delete useless routes from the routing table of the neighbor and also avoid route loops.

Figure 16-12 Schematic diagram of poison reverse



As shown in Figure 16-12, if poison reverse is not configured, Router B sends Router A a route that is learned from Router A. The cost of the route from Router A to network 10.0.0.0 is 1. If the route from Router A to network 10.0.0.0 is unreachable and Router B keeps sending Router A routes to network 10.0.0.0 because Router B failed to receive the route update packet from Router , forming a a route loop.

With poison reverse configured, if Router A sends Router B a message that the route received from Router B is unreachable, Router B does not learns the unreachable route from Router A, which avoids route loops.

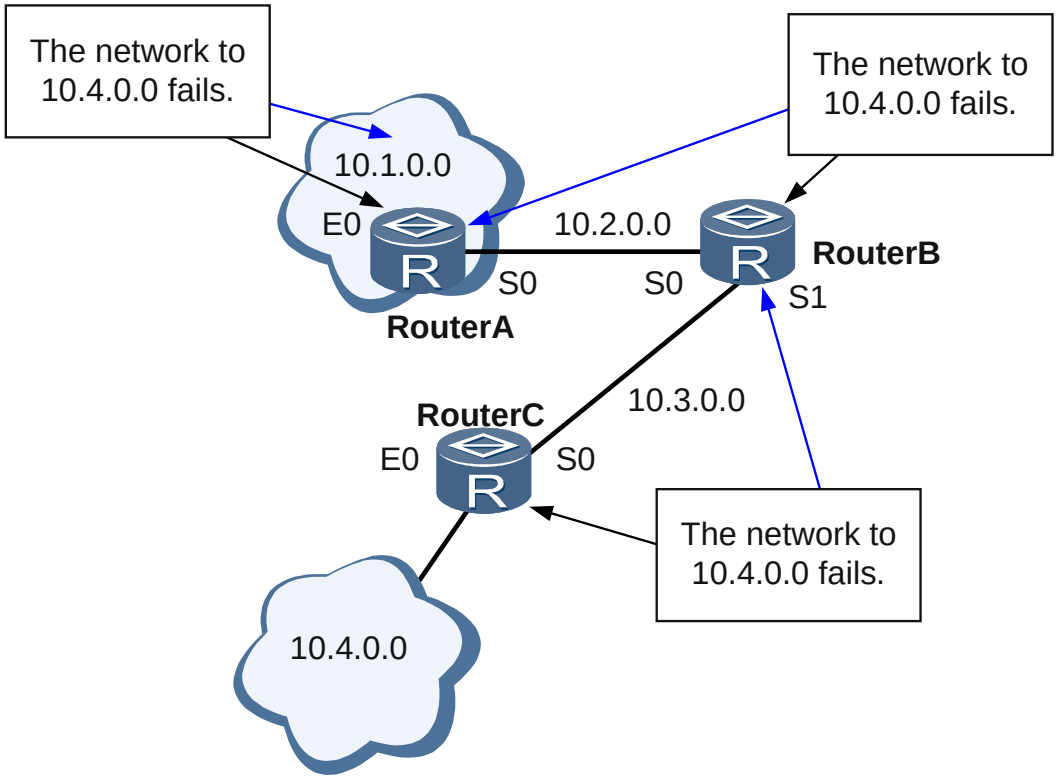
If both poison reverse and split horizon are configured, split horizon (the route learned from an interface is not sent back through the interface) will be replaced by poison reverse.

16.6.7 Triggered Update

Triggered update occurs when the local routing information changes and the local router immediately notifies its neighbors of the changes in routing information by sending the triggered update packet.

Triggered update shortens the network convergence time. When the local routing information changes, the local router immediately notifies its neighbors of the changes in routing information rather than wait for a periodic update.

Figure 16-13 Triggered update



In the example in Figure 16-13, when network 10.4.0.0 is unreachable, Router C learns the information first. Usually, the route update message is sent to neighbors every 30s. If the update message of Router B is sent to Router C when Router C is waiting for the route update message, Router C learns the faulty route to network 10.4.0.0 from Router B. In this case, the routes from Router B or Router C to network 10.4.0.0 point to Router C or Router B respectively, which forms a route loop. If Router C detects a network fault and immediately sends a route update message to Router B before the new update interval reaches Router B. The routing table of Router B is updated in time, and routing loops are avoided.

Another scenario that triggers updates is when the next hop of the route is unavailable because the link is faulty. The local device needs to notify neighboring device about the routes'unreachability. The local device sets the cost the route to 16 and advertising the route. This is also called *route-withdrawal*.

16.6.8 Route Aggregation

When different subnet routes in the same natural network segment are transmitted to other network segments, these routes are aggregated into one route of the same segment. This process is called route aggregation. RIP-1 packets do not carry mask information, so RIP-1 can advertise only the routes with natural masks. Because RIP-2 packets do carry mask information, RIP-2 supports subnetting.

RIP-2 route convergence can improve extensibility and efficiency and minimize the routing table of a large-scale network.

Route convergence is classified into two types as follows:

- Classful aggregation based on RIP processes:

  For example, router 10.1.1.0/24 (metric=2) and router 10.1.2.0/24 (metric=3) are aggregated as an aggregated route (10.0.0.0/8(metric=2)) in the natural network segment. Because RIP-2 aggregation is classful, obtains the optimal metric.
- Interface-based aggregation:

  A user can specify an aggregation address.

  For example, router 10.1.1.0/24(metric=2) and router 10.1.2.0/24 (metric=3) are aggregated as an aggregated route (10.1.0.0/16(metric=2)).

16.6.9 Multi-process and Multi-instance

For easy management and effective control, RIP supports the features multi-process and multi-instance. The multi-process feature allows a set of interfaces to be associated with a specific RIP process. This ensures that the process performs all the protocol operations only on this set of interfaces. Therefore, multiple RIP processes can work on a single router and each process is responsible for a unique set of interfaces. In addition, the routing data is independent between RIP processes. However, routes can be imported between processes.

For routers that support VPN, you can associate each RIP process with a specific VPN instance. In this case, all the interfaces attached to the RIP process should be associated with the RIP-process-related VPN instance.

16.6.10 Hot Backup

Routers with distributed architecture support the RIP Hot Standby (HSB) feature. RIP backs up data from the Active Main Board (AMB) to the Standby Main Board (SMB). Whenever the AMB fails, the SMB becomes active. In this manner, RIP, being free from active/standby switchover, proceeds to work normally.

RIP supports only the backup of RIP configurations. RIP performs Graceful Restart (GR) to resend a routing request to neighbors and synchronize route database.

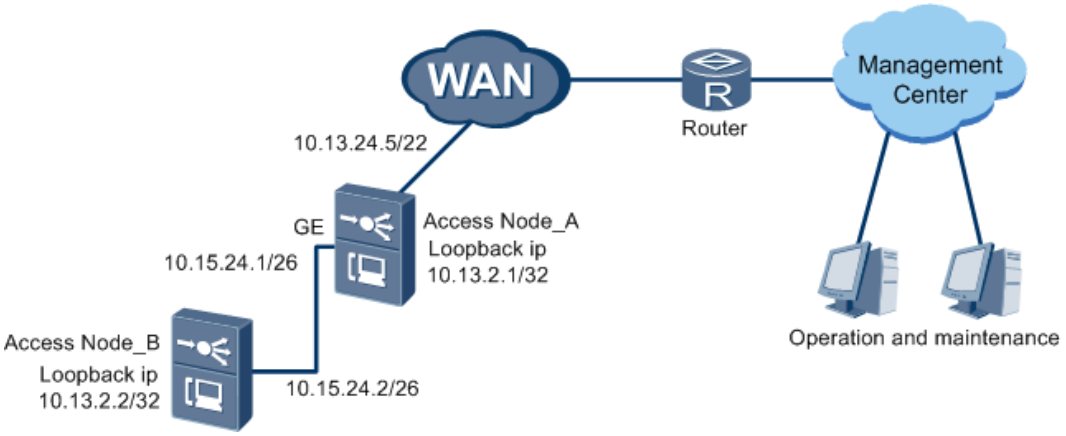
16.6.11 Configuration Example of RIP

This topic provides an example for configuring RIP on the MA5600T/MA5603T/MA5608T.

Service Requirements

- MA5600T/MA5603T/MA5608T_A is subtended with MA5600T/MA5603T/MA5608T_B through port 0/19/1, and uses port 0/19/0 to transmit services in the upstream. Besides, it connects to the management center network through the WAN.
- RIP is enabled on MA5600T/MA5603T/MA5608T_A and MA5600T/MA5603T/MA5608T_B so that the administrator can access MA5600T/MA5603T/MA5608T_A and MA5600T/MA5603T/MA5608T_B through the RIP route. Then, you can operate and maintain MA5600T/MA5603T/MA5608T_A and MA5600T/MA5603T/MA5608T_B.

Figure 16-14 Example network for configuring RIP



Data Plan

Table 16-6 provides the data plan for configuring RIP.

Table 16-6 Data plan for configuring RIP

Item	Data
MA5600T/MA5603T/MA5608T_A	Upstream port: 0/19/0 Administration VLAN: smart VLAN 100 IP address of the Layer 3 interface in the administration VLAN: 10.13.24.5/22 Loopback interface address: 10.13.2.1/32
	RIP version: V2 RIP route filtering policy: filtering routes based on the IP address prefix list "abc". Only the routes with the IP addresses 10.13.2.1 and 10.13.2.2 can be advertised through the Layer 3 interface of VLAN 100.
	Subtending port: 0/19/0 Subtending administration VLAN: smart VLAN 10 IP address of the Layer 3 interface in the subtending administration VLAN: 10.15.24.1/26
MA5600T/MA5603T/MA5608T_B	Subtending port: 0/19/1 Administration VLAN: smart VLAN 10 IP address of the Layer 3 interface in the administration VLAN: 10.15.24.2/26 Loopback interface address: 10.13.2.2/32
	RIP version: V2 RIP route filtering policy: filtering routes based on the IP address prefix list "abc". Only the route with the IP address 10.13.2.2 can be advertised through the Layer 3 interface of VLAN 10.

Procedure

- Configure MA5600T/MA5603T/MA5608T_A.
  a. Configure the RIP-supported Layer 3 interface.

```
huawei(config)#vlan 100 smart
huawei(config)#port vlan 100 0/19 0
huawei(config)#interface vlanif 100
huawei(config-if-vlanif100)#ip address 10.13.24.5 22
huawei(config-if-vlanif100)#quit
huawei(config)#interface loopBack 0
huawei(config-if-loopback0)#ip address 10.13.2.1 32
huawei(config-if-loopback0)#quit
```

  b. Enable RIP.

```
huawei(config)#rip 1
huawei(config-rip-1)#network 10.13.24.0
huawei(config-rip-1)#network 10.13.2.0
huawei(config-rip-1)#version 2
huawei(config-rip-1)#quit
```

  c. Configure the route filtering policy.

```
huawei(config)#ip ip-prefix abc permit 10.13.2.1 32
huawei(config)#ip ip-prefix abc permit 10.13.2.2 32
huawei(config)#rip 1
huawei(config-rip-1)#filter-policy ip-prefix abc export vlanif 100
huawei(config-rip-1)#quit
```

  d. Configure the subtending port.

```
huawei(config)#vlan 10 smart
huawei(config)#port vlan 10 0/19 1
huawei(config)#interface giu 0/19
huawei(config-if-giu-0/19)#network-role 1 cascade
huawei(config-if-giu-0/19)#quit
huawei(config)#interface vlanif 10
huawei(config-if-vlanif10)#ip address 10.15.24.1 26
huawei(config-if-vlanif10)#quit
```

  e. Enable RIP on the subtending port.

```
huawei(config)#rip 1
huawei(config-rip-1)#network 10.15.24.0
huawei(config-rip-1)#quit
```

  f. Save the data.

```
huawei(config)#save
```

- Configure MA5600T/MA5603T/MA5608T_B.
  a. Configure the RIP-supported Layer 3 interface.

```
huawei(config)#vlan 10 smart
huawei(config)#port vlan 10 0/19 0
huawei(config)#interface vlanif 10
```

```
huawei(config-if-vlanif10)#ip address 10.15.24.2 26
huawei(config-if-vlanif10)#quit
huawei(config)#interface loopBack 0
huawei(config-if-loopback0)#ip address 10.13.2.2 32
huawei(config-if-loopback0)#quit
```

b. Enable RIP.

```
huawei(config)#rip 1
huawei(config-rip-1)#network 10.15.24.0
huawei(config-rip-1)#network 10.13.2.0
huawei(config-rip-1)#version 2
huawei(config-rip-1)#quit
```

c. Configure the route filtering policy.

```
huawei(config)#ip ip-prefix abc permit 10.13.2.2 32
huawei(config)#rip 1
huawei(config-rip-1)#filter-policy ip-prefix abc export vlanif 10
huawei(config-rip-1)#quit
```

d. Save the data.

```
huawei(config)#save
```

----End

Result

The maintenance terminal of the administration center can access MA5600T/MA5603T/MA5608T_A and MA5600T/MA5603T/MA5608T_B, and operate and maintain the two devices.

Configuration File

Configuration on MA5600T/MA5603T/MA5608T_A

```
vlan 100 smart
port vlan 100 0/19 0
interface vlanif 100
ip address 10.13.24.5 22
quit
interface loopBack 0
ip address 10.13.2.1 32
quit
rip 1
network 10.13.24.0
network 10.13.2.0
version 2
quit
ip ip-prefix abc permit 10.13.2.1 32
ip ip-prefix abc permit 10.13.2.2 32
rip 1
filter-policy ip-prefix abc export vlanif 100
quit
```

```
vlan 10 smart
port vlan 10 0/19 1
interface giu 0/19
network-role 1 cascade
quit
interface vlanif 10
ip address 10.15.24.1 26
quit
rip 1
network 10.15.24.0
quit
save
```

Configuration on MA5600T/MA5603T/MA5608T_B

```
vlan 10 smart
port vlan 10 0/19 0
interface vlanif 10
ip address 10.15.24.2 26
quit
interface loopBack 0
ip address 10.13.2.2 32
quit
rip 1
network 10.15.24.0
network 10.13.2.0
version 2
quit
ip ip-prefix abc permit 10.13.2.2 32
rip 1
filter-policy ip-prefix abc export vlanif 10
quit
save
```

16.6.12 References

The following table lists the references that apply in this chapter.

Document No.	Document Name	Protocol Compliance
RFC 1058	Routing Information Protocol	Fully compliant.
RFC 2453	RIP Version 2	Fully compliant.

16.7 RIPng

16.7.1 Introduction to RIPng

Definition

RIPng is an IPv6 extension of RIP-2 on the original IPv4 network. Most RIP concepts can be applied to RIPng.

RIPng, based on the Distance Vector (D-V) algorithm, is a routing protocol that measures the distance (metrics or cost) to the destination host by Hop Count (HC). According to RIPng, the HC from a router to its directly connected network is 0, and the HC from a router to a network that is reachable through another router is 1, and so on. When the HC reaches 16, the destination network or host is defined as unreachable.

For adaption to the IPv6 network, RIPng is derived from RIP with changes as follows:

- UDP port number: RIPng uses UDP port number 521 to send and receive routing information.
- Multicast address: RIPng uses FF02::9 as the multicast address of a RIPng router in the local scope of the links.
- Prefix length: RIPng uses a 128-bit (the mask length) prefix in the destination address.
- Next hop address: RIPng uses a 128-bit IPv6 address.
- Source address: RIPng uses the local link address FE80::/10 as the source address to send RIPng update packets.

Purpose

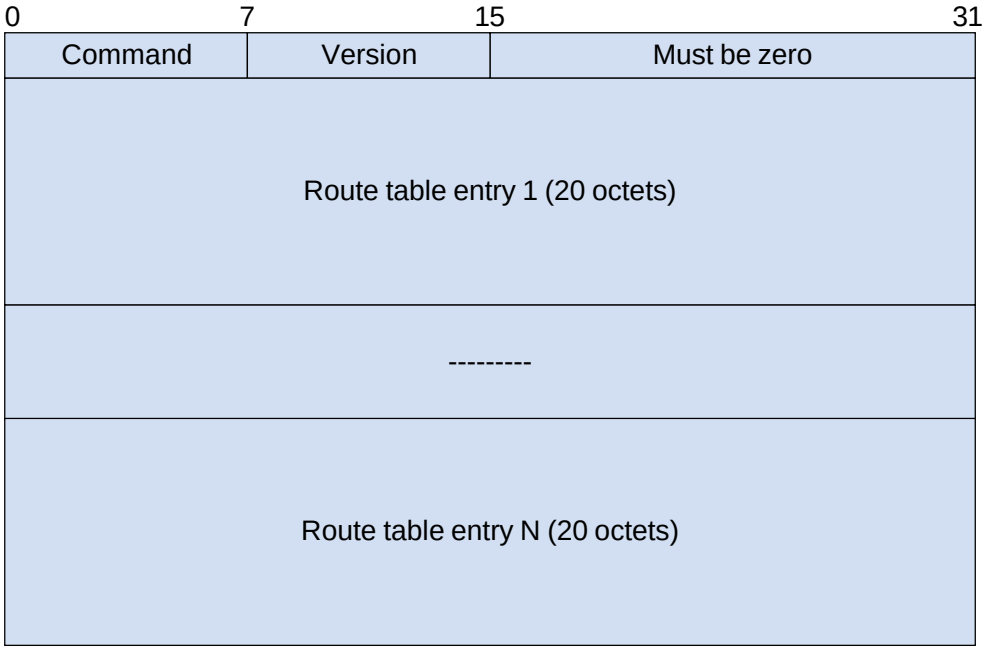
RIPng is developed by extending RIP to support IPv6.

16.7.2 RIPng Packet Format

A RIPng packet is composed of a header and several route table entries (RTEs). In a RIPng packet, the maximum number of RTEs is determined by the MTU value of the interface.

Figure 16-15 shows the basic format of a RIPng packet.

Figure 16-15 RIPng packet format



A RIPng packet contains two types of RTEs as follows:

- Next hop RTE: is located before the IPv6-prefix RTEs that have the same next hop. It defines the IPv6 address of the next hop.
- IPv6-prefix RTE: is located after a next-hop RTE. Several different IPv6-prefix RTEs can exist after the next-hop RTE. It describes the destination IPv6 address and the cost in the RIPng routing table.

Figure 16-16 shows the format of the next-hop RTE.

Figure 16-16 Format of the next hop RTE

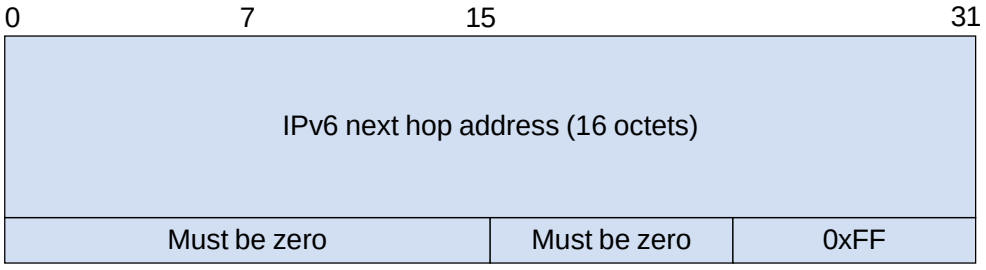
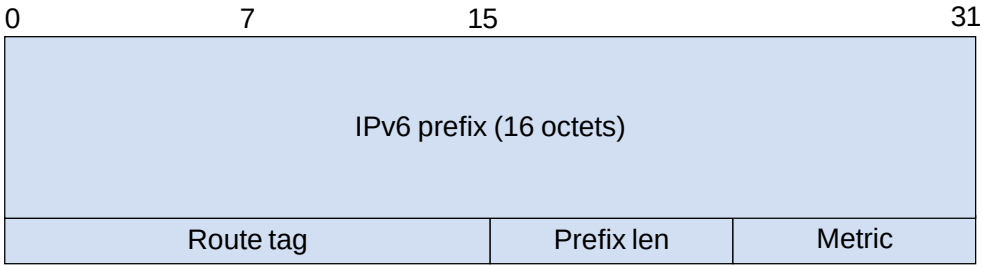


Figure 16-17 shows the format of the IPv6-prefix RTE.

Figure 16-17 Format of the IPv6-prefix RTE



16.7.3 Timer

RIPng uses the following three timers:

- Update timer: The timer triggers the sending of update packets every 30s. This timer synchronizes RIPng routes on the network.
- Age timer: If a RIPng router does not receive any update packet from its neighbors in the aging time, the RIPng router considers the route to its neighbors unreachable.
- Garbage-Collect timer: If the route is no longer valid after the timer times out, the entry is removed from the RIPng routing table.

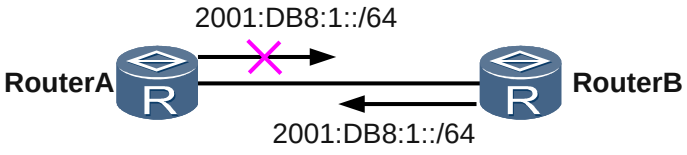
The following describes the relationship among the three timers:

The advertisement of RIPng routing update is triggered by the update timer every 30 seconds. Each entry is associated with two timers, the age timer and the garbage-collect timer. When a route is learned and installed in the routing table, the age timer is initialized. If no Update packet is received from the neighbor for 180 seconds, the metric of the route is set to 16. At the same time, the garbage-collect timer is initialized. If no Update packet is received for 120 seconds, the entry is deleted after the garbage-collect timer times out.

16.7.4 Split Horizon

The principle of split horizon is that a route learnt by RIPng on an interface is not sent to neighbors from the interface. This reduces bandwidth consumption and avoids route loops.

Figure 16-18 Schematic diagram of split horizon



As shown in Figure 16-18, Router B sends a route to network 2001:DB8:1::/64 to Router A and Router A does not send the route back to Router B.

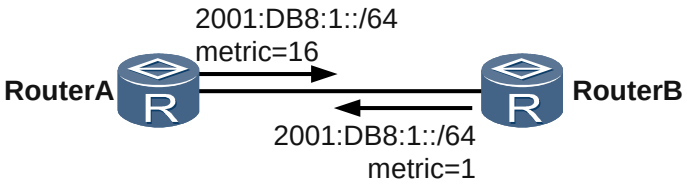
16.7.5 Poison Reverse

The principle of poison reverse is that RIPng sets the cost of the route learnt from an interface of a neighbor to 16 (specifying the route as unreachable) and then sends the route from the

interface to the neighbor. In this way, RIPng can delete useless routes from the routing table of the neighbor.

Poison reverse of RIPng can also avoid route loops.

Figure 16-19 Schematic diagram of poison reverse



As shown in Figure 16-19, if poison reverse is not configured, Router B sends Router A a route that is learnt from Router A. The cost of the route from Router A to network 2001:DB8:1::/64 is 1. When the route from Router A to network 2001:DB8:1::/64 becomes unreachable and Router B does not receive the update packet from Router A and therefore keeps sending Router A the route from Router A to network 2001:DB8:1::/64, a route loop occurs.

If Router A sends Router B a message that the route is unreachable after receiving a route from Router B, Router B no longer learns the reachable route from Router A, avoiding route loops.

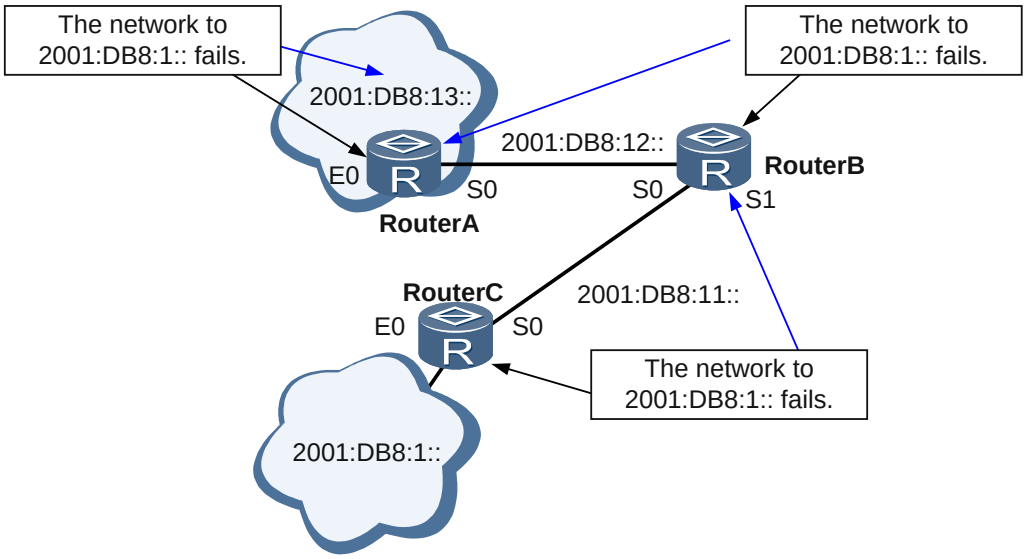
If both poison reverse and split horizon are configured, simple split horizon (the route learnt from an interface is not sent back through the interface) is replaced by poison reverse.

16.7.6 Triggered Update

Triggered update occurs when local routing information changes and then the local router immediately notifies its neighbors of the changes of routing information by sending the triggered update packet.

Triggered update shortens the network convergence time. When local routing information changes, the local router immediately notifies its neighbors of the changes of routing information rather than waiting for periodical update.

Figure 16-20 Schematic diagram of triggered update



As shown in Figure 16-20, when network 2001:DB8:1:: is unreachable, Router C learns the information first. Usually, the route update message is periodically sent to neighbors. For example, RIPng sends the route update message every 30s. If the update message of Router B is sent to Router C when Router C is waiting for the route update message, Router C learns the faulty route to network 2001:DB8:1:: from Router B. In this case, the routes from Router B or Router C to network 2001:DB8:1:: point to Router C or Router B respectively, forming a route loop. If Router C detects a network fault and immediately sends a route update message to Router B before the new update interval reaches. Consequently, the routing table of Router B is updated in time, and routing loops are avoided.

There is another mode of triggering updates: The next hop of the route is unavailable because the link is faulty. The local Router needs to notify neighboring Router about the unreachability of this route. This is done by setting the cost of the route as 16 and advertising the route. This is also called route-withdrawal.

16.7.7 Route Aggregation

RIPng route aggregation is implemented by aggregating all routes advertised on an interface according to the longest match rule.

RIPng route aggregation can improve extensibility and efficiency and minimize the routing table of a large-scale network.

Implementation of route aggregation:

For example, RIPng advertises two routes, 2001:DB8:11::24 Metric=2 and 2001:DB8:12::34 Metric=3, from an interface, and the aggregation route configured on the interface is 2001:DB8::/32. In this manner, the finally advertised route is 2001:DB8::/32 Metric=2.

16.7.8 Multi-process

For easy management and effective control, RIPng supports multi-process and multi-instance. The multi-process feature allows a set of interfaces to be associated with a specific RIPng process. This ensures that the specific RIPng process performs all the protocol operations only on this set of interfaces. Therefore, multiple RIPng processes can work on a single router and

each process is responsible for a unique set of interfaces. In addition, the routing data is independent between RIPng processes; however, routes can be imported between processes.

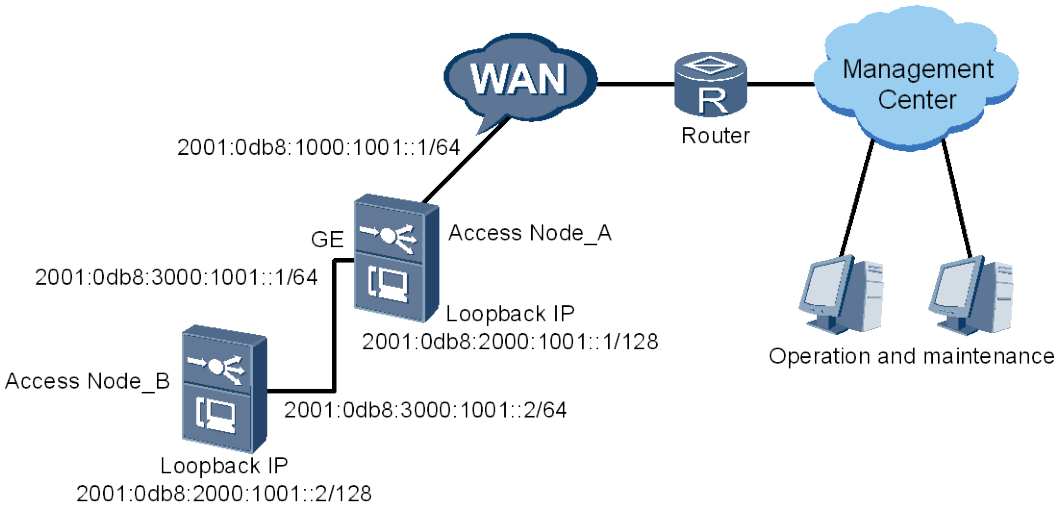
16.7.9 Configuration Example of RIPng

This topic provides an example for configuring RIPng on the MA5600T/MA5603T/MA5608T.

Service Requirements

- Access node_A is subtended with Access node_B through port 0/19/1, and uses port 0/19/0 to transmit services in the upstream. Besides, it connects to the management center network through the WAN.
- RIPng is enabled on Access node_A and Access node_B so that the administrator can access Access node_A and Access node_B through the RIPng route. Then, you can operate and maintain Access node_A and Access node_B.

Figure 16-21 Example network for configuring RIPng



Data Plan

Table 16-7 provides the data plan for configuring RIPng.

Table 16-7 Data plan for configuring RIPng

Item	Data
Access node_A	Upstream port: 0/19/0 Administration VLAN: smart VLAN 100 IPv6 address of the Layer 3 interface in the administration VLAN: 2001:0db8:1000:1001::1/64 Loopback interface address: 2001:0db8:2000:1001::1/128
	RIPng route filtering policy: filtering routes based on the IP address prefix list "abc". Only the routes with the IP addresses 2001:0db8:2000:1001::1 and 2001:0db8:2000:1001::2 can be

Item	Data
	advertised through the Layer 3 interface of VLAN 100.
	Subtending port: 0/19/0 Subtending administration VLAN: smart VLAN 10 IPv6 address of the Layer 3 interface in the subtending administration VLAN: 2001:0db8:3000:1001::1/64
Access node_B	Subtending port: 0/19/1 Administration VLAN: smart VLAN 10 IPv6 address of the Layer 3 interface in the administration VLAN: 2001:0db8:3000:1001::2/64 Loopback interface address: 2001:0db8:2000:1001::2/128
	RIPng route filtering policy: filtering routes based on the IP address prefix list "abc". Only the route with the IP address 2001:0db8:2000:1001::2 can be advertised through the Layer 3 interface of VLAN 10.

Procedure

- Configure Access node_A.
  a. Configure the RIPng-supported Layer 3 interface.

```
huawei(config)#vlan 100 smart
huawei(config)#port vlan 100 0/19 0
huawei(config)#ipv6
huawei(config)#interface vlanif 100
huawei(config-if-vlanif100)#ipv6 enable
huawei(config-if-vlanif100)#ipv6 address 2001:0db8:1000:1001::1 64
huawei(config-if-vlanif100)#quit
huawei(config)#interface loopBack 0
huawei(config-if-loopback0)#ipv6 enable
huawei(config-if-loopback0)#ipv6 address 2001:0db8:2000:1001::1/128
huawei(config-if-loopback0)#quit
```

  b. Enable RIPng.

```
huawei(config)#ripng 1
huawei(config-ripng-1)#quit
huawei(config)#interface vlanif 100
huawei(config-if-vlanif100)#ripng 1 enable
huawei(config-if-vlanif100)#quit
huawei(config)#interface loopBack 0
huawei(config-if-loopback0)#ripng 1 enable
huawei(config-if-loopback0)#quit
```

  c. Configure the route filtering policy.

```
huawei(config)#ip ipv6-prefix abc permit 2001:0db8:2000:1001::1 128
huawei(config)#ip ipv6-prefix abc permit 2001:0db8:2000:1001::2 128
huawei(config)#ripng 1
```

```
huawei(config-ripng-1)#filter-policy ipv6-prefix abc export
huawei(config-ripng-1)#quit
```

d. Configure the subtending port.

```
huawei(config)#vlan 10 smart
huawei(config)#port vlan 10 0/19 1
huawei(config)#interface giu 0/19
huawei(config-if-giu-0/19)#network-role 1 cascade
huawei(config-if-giu-0/19)#quit
huawei(config)#interface vlanif 10
huawei(config-if-vlanif10)#ipv6 enable
huawei(config-if-vlanif10)#ipv6 address 2001:0db8:3000:1001::1 64
huawei(config-if-vlanif10)#quit
```

e. Enable RIPng on the subtending port.

```
huawei(config)#interface vlanif 10
huawei(config-if-vlanif10)#ripng 1 enable
huawei(config-if-vlanif10)#quit
```

f. Save the data.

```
huawei(config)#save
```

- Configure Access node_B.

a. Configure the RIPng-supported Layer 3 interface.

```
huawei(config)#vlan 10 smart
huawei(config)#port vlan 10 0/19 0
huawei(config)#ipv6
huawei(config)#interface vlanif 10
huawei(config-if-vlanif10)#ipv6 enable
huawei(config-if-vlanif10)#ipv6 address 2001:0db8:3000:1001::2 64
huawei(config-if-vlanif10)#quit
huawei(config)#interface loopback 0
huawei(config-if-loopback0)#ipv6 enable
huawei(config-if-loopback0)#ipv6 address 2001:0db8:2000:1001::2 128
huawei(config-if-loopback0)#quit
```

b. Enable RIPng.

```
huawei(config)#ripng 1
huawei(config-ripng-1)#quit
huawei(config)#interface vlanif 10
huawei(config-if-vlanif10)#ripng 1 enable
huawei(config-if-vlanif10)#quit
huawei(config)#interface loopback 0
huawei(config-if-loopback0)#ripng 1 enable
huawei(config-if-loopback0)#quit
```

c. Configure the route filtering policy.

```
huawei(config)#ip ipv6-prefix abc permit 2001:0db8:2000:1001::2
huawei(config)#ripng 1
huawei(config-ripng-1)#filter-policy ipv6-prefix abc export
huawei(config-ripng-1)#quit
```

d. Save the data.

```
huawei(config)#save
```

----End

Result

The maintenance terminal of the administration center can access Access node_A and Access node_B, and operate and maintain the two devices.

Configuration File

Configuration on Access node_A

```
vlan 100 smart
port vlan 100 0/19 0
ipv6
interface vlanif 100
ipv6 enable
ipv6 address 2001:0db8:1000:1001::1 64
quit
interface loopBack 0
ipv6 enable
ipv6 address 2001:0db8:2000:1001::1/128
quit
ripng 1
quit
interface vlanif 100
ripng 1 enable
quit
interface loopBack 0
ripng 1 enable
quit
ip ipv6-prefix abc permit 2001:0db8:2000:1001::1 128
ip ipv6-prefix abc permit 2001:0db8:2000:1001::2 128
ripng 1
filter-policy ipv6-prefix abc export
quit
vlan 10 smart
port vlan 10 0/19 1
interface giu 0/19
network-role 1 cascade
quit
interface vlanif 10
ipv6 enable
ipv6 address 2001:0db8:3000:1001::1 64
ripng 1 enable
quit
save
```

Configuration on Access node_B

```
vlan 10 smart
port vlan 10 0/19 0
interface vlanif 10
```

```
ipv6
interface vlanif 10
ipv6 enable
ipv6 address 2001:0db8:3000:1001::2 64
quit
interface loopback 0
ipv6 enable
ipv6 address 2001:0db8:2000:1001::2 128
quit
ripng 1
quit
interface vlanif 10
ripng 1 enable
quit
interface loopback 0
ripng 1 enable
quit
ip ipv6-prefix abc permit 2001:0db8:2000:1001::2
ripng 1
filter-policy ipv6-prefix abc export
quit
save
```

16.7.10 References

The following table lists the references.

Document No.	Document Name	Protocol Compliance
RFC 2080	RIPng for IPv6	Fully compliant.

16.8 IS-IS

Intermediate System-to-Intermediate System (IS-IS) is a link state protocol. It uses the shortest path first (SPF) algorithm to calculate routes. IS-IS is one of Interior Gateway Protocols and is used inside of an autonomous system.

16.8.1 Introduction to IS-IS

Definition

Intermediate System-to-Intermediate System (IS-IS) is a dynamic routing protocol initially designed by the International Organization for Standardization (ISO) for its Connectionless Network Protocol (CLNP).

To support IP routing, the Internet Engineering Task Force (IETF) extends and modifies IS-IS in RFC 1195. This modification enables IS-IS to be applied to TCP/IP and OSI environments. This type of IS-IS is called Integrated IS-IS or Dual IS-IS.

The term IS-IS used in this document refers to Integrated IS-IS, unless otherwise stated.

Purpose

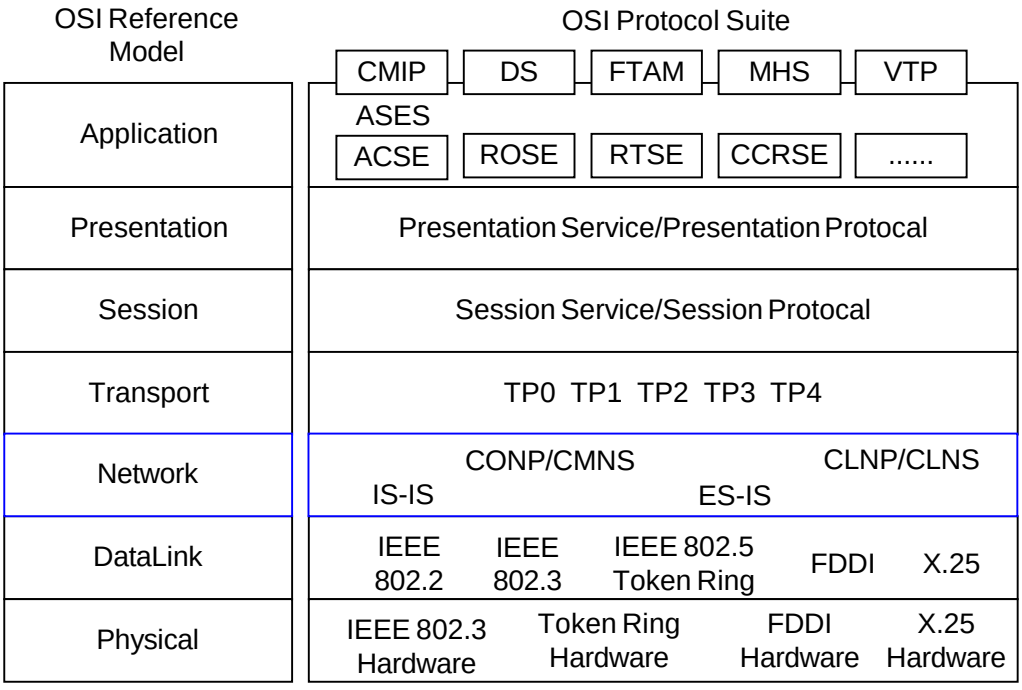
As an Interior Gateway Protocol (IGP), IS-IS is used in Autonomous Systems (ASs). IS-IS is a link state protocol. It uses the Shortest Path First (SPF) algorithm to calculate routes.

16.8.2 Basic Concepts of IS-IS

Development of IS-IS

CLNP is a Layer 3 protocol in the OSI model posed by the ISO. IS-IS was initially designed by the ISO and is used as a routing protocol based on CLNP addressing.

Figure 16-22 OSI model



OSI adopts systemized (or hierarchical) addressing. The services on the transport layer in OSI can be addressed through the Network Service Access Point (NSAP).

OSI uses the following terms:

- CLNS: indicates the Connectionless Network Service.
- CLNP: indicates the Connectionless Network Protocol.
- CMNS: indicates the Connection-Mode Network Service.
- CONP: indicates the Connection-Oriented Network Protocol.

OSI implements CLNS through CLNP, and implements CMNS through CONP.

CLNS is implemented through the following protocols:

- CLNP: is similar to the IP protocol in TCP/IP.
- IS-IS: is the routing protocol of an intermediate system.
- ES-IS: is the protocol used between a host system and an intermediate system. It is similar to ARP or ICMP in IP.

Table 16-8 Comparison of concepts in OSI and IP

Abbreviation	Concepts in OSI	Concepts in IP
IS	Intermediate System	Router
ES	End System	Host
DIS	Designated Intermediate System	Designated Router (DR) in OSPF
SysID	System ID	Router ID in OSPF
PDU	Protocol Data Unit	IP packet
LSP	Link State Protocol data unit	OSPF LSA
NSAP	Network Service Access Point	IP address

With the popularity of TCP/IP, the IETF extends and modifies IS-IS in RFC 1195 to support IP routing. This modification enables IS-IS to be applied to TCP/IP and OSI environments. This type of IS-IS is called Integrated IS-IS or Dual IS-IS.

Address Structure of IS-IS

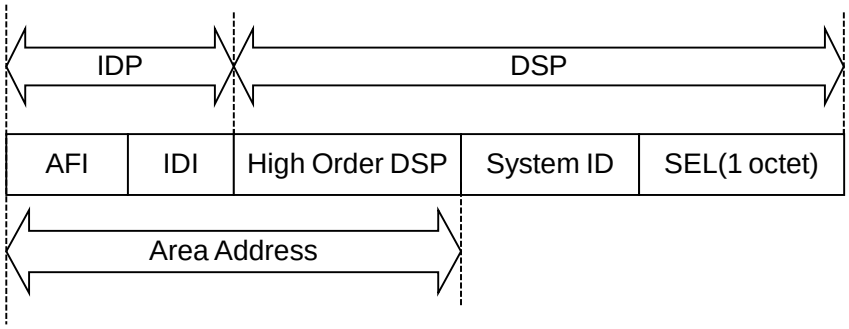
In OSI, the NSAP is an address used to locate resources. The ISO has adopted the NSAP address structure shown in Figure 16-23. NSAP is composed of the Initial Domain Part (IDP) and the Domain Specific Part (DSP). The IDP is equal to the network ID in an IP address, and DSP is equal to the subnet number and host address in an IP address.

As defined by the ISO, the IDP consists of the Authority and Format Identifier (AFI) and the Initial Domain Identifier (IDI). The AFI specifies the address assignment mechanism and address format; the IDI identifies a domain.

The DSP consists of the High Order DSP (HODSP), system ID, and NSAP Selector (SEL). The HODSP is used to divide areas, the system ID identifies a host, and the SEL indicates the service type.

The lengths of the IDP and the DSP are variable. The maximum length of the NSAP is 20 bytes and its minimum length is 8 bytes.

Figure 16-23 Networking for IS-IS address structure



The components in the address structure are described as follows:

- Area address

  Together with the HODSP of the DSP, the IDP can identify a routing domain and the areas in a routing domain. The combination of the IDP and HODSP is referred to as an *area address*, which is equal to an area number in OSPF. An area address is used to uniquely identify the area in the routing domain. The area addresses of routers in the same Level-1 area must be the same, while the area addresses of routers in the Level-2 area can be different.

  In general, a router can be configured with only one area address. The area address of all nodes in an area must be the same. In the implementation of a device, an IS-IS process can be configured with a maximum of three area addresses to support seamless combination, division, and transformation of areas.

- System ID

  A system ID uniquely identifies a host or a router in an area. In the device, the fixed length of the system ID is 48 bits (6 bytes).

  In actual applications, a router ID corresponds to a system ID. If a router takes the IP address 168.10.1.1 of Loopback 0 as its router ID, its system ID used in IS-IS can be obtained in the following way:

  - To extend each part of the IP address 168.10.1.1 to 3 bits, add 0 to the front of any part that is shorter than 3 bits.
  - Divide the extended address 168.010.001.001 into three parts, with each part consisting of four decimal digits.
  - The reconstructed address 1680.1000.1001 is the system ID.

  You can specify a system ID in many ways. You need to ensure that the system ID uniquely identifies a host or a router.

- SEL

  The role of an SEL (also referred to as NSAP Selector or N-SEL) is similar to that of the "protocol identifier" of IP. A transport protocol matches an SEL. The SEL is always "00" in IP.

- NET

  A Network Entity Title (NET) indicates the network layer information of an IS itself. It does not contain the transport layer information (SEL = 0). A NET can be regarded as a special NSAP. The length of the NET field is the same as that of an NSAP. Its maximum length is 20 bytes and its minimum length is 8 bytes. When configuring IS-IS on a router, you can configure only a NET instead of an NSAP.

In general, an IS-IS process is configured with only one NET. When an area needs to be redefined, such as being combined with other areas or divided into sub-areas, you can configure the router with multiple NETs to ensure the correctness of routes.

You can configure an IS-IS process can be configured with a maximum of three area addresses, and a maximum of three NETs can be configured. When you configure multiple NETs, ensure that their system IDs are the same.

For example, in the NET ab.cdef.1234.5678.9abc.00 the area is ab.cdef, the system ID is 1234.5678.9abc, and the SEL is 00.

NOTE

The routers in an area must have the same area address.

IS-IS PDU Format

The types of PDUs for IS-IS include Hello, LSPs, CSNPs, and PSNPs.

Table 16-9 PDU types

Type Value	PDU Type	Name
15	Level-1 LAN IS-IS Hello PDU	L1 LAN IIH
16	Level-2 LAN IS-IS Hello PDU	L2 LAN IIH
17	Point-to-Point IS-IS Hello PDU	P2P IIH
18	Level-1 Link State PDU	L1 LSP
20	Level-2 Link State PDU	L2 LSP
24	Level-1 Complete Sequence Numbers PDU	L1 CSNP
25	Level-2 Complete Sequence Numbers PDU	L2 CSNP
26	Level-1 Partial Sequence Numbers PDU	L1 PSNP
27	Level-2 Partial Sequence Numbers PDU	L2 PSNP

- Hello packet format

  Hello packets, also called the IS-to-IS Hello PDUs (IIH), are used to set up and maintain neighbor relationships. Among them, Level-1 LAN IIHs are applied to the Level-1 routers on broadcast LANs; Level-2 LAN IIHs are applied to the Level-2 routers on broadcast LANs; and P2P IIHs are applied to non-broadcast networks. Hello packets in different networks have different formats.

  Figure 16-24 shows the format of a Hello packet in a broadcast network The area highlighted in blue is the common header.

Figure 16-24 Level-1 or Level-2 LAN IIH Hello packet

	No. of Octets
Intradomain Routeing Protocol Discriminator	1
Length Indicator	1
Version/Protocol ID Extension	1
ID Length	1
R R R PDU Type	1
Version	1
Reserved	1
Maximum Area Address	1
Reserved/Circuit Type	1
Source ID	ID Length
Holding Time	2
PDU Length	2
R Priority	1
LAN ID	ID Length+1
Variable Length Fields	

Figure 16-25 shows the format of a Hello packet in a P2P network.

Figure 16-25 P2P IIH Hello packet

	No. of Octets
Intradomain Routeing Protocol Discriminator	1
Length Indicator	1
Version/Protocol ID Extension	1
ID Length	1
R R R PDU Type	1
Version	1
Reserved	1
Maximum Area Address	1
Reserved/Circuit Type	1
Source ID	ID Length
Holding Time	2
PDU Length	2
Local Circuit ID	1
Variable Length Fields	

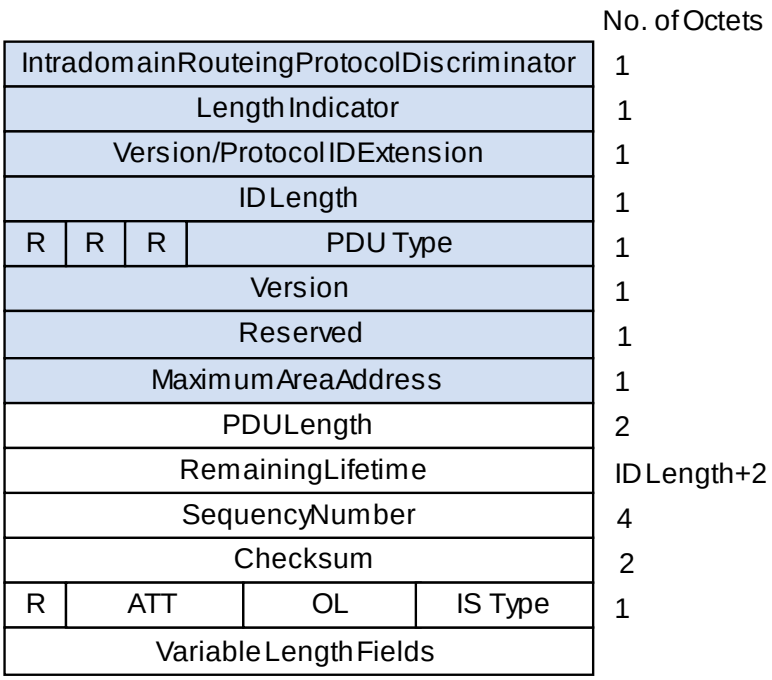
As shown in Figure 16-25, most fields in a P2P IIH are the same as those in a LAN IIH. The P2P IIH does not have the priority and LAN ID fields, but has a local circuit ID field. The local circuit ID indicates the local link ID.

- LSP packet format

  Link State PDUs (LSPs) are used to exchange link-state information. There are two types of LSPs: Level-1 and Level-2. Level-1 IS-IS transmits Level-1 LSPs; Level-2 IS-IS transmits Level-2 LSPs; and Level-1-2 IS-IS can transmit both Level-1 and Level-2 LSPs.

  Level-1 and Level-2 LSPs have the same format, as shown in Figure 16-26.

Figure 16-26 Level-1 or Level-2 LSP packet



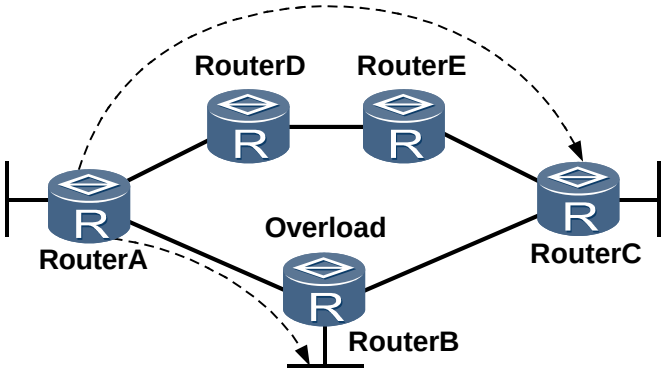
The main fields of the level-1 or Level-2 LSP packet are described as follows:

- OL: indicates LSDB overload.

  LSPs with the overload bit are still flooded on the network, but the LSPs are not used when routes that pass through a router configured with the overload bit are calculated. After a router is configured with the overload bit, other routers ignore the router when performing the SPF calculation. Only the direct routes of the router are considered.

  As shown in Figure 16-27, packets from Router A to Router C are all forwarded by Router B. If the OL field is set to 1 on Router B, however, Router A detects that the LSDB of Router B is incomplete. Router A then forwards the packets to Router C through Router D and Router E, but the packets to the destination that is directly connected to Router B are forwarded normally.

Figure 16-27 Networking for LSDB overload



- IS Type: indicates the type of IS-IS that generates the LSP.

  The IS Type specifies whether the level of IS-IS is Level-1 or Level-2 (01 indicates Level-1; 11 indicates Level-2).

- **SNP Format**

  Sequence Number PDUs (SNPs) describe the LSPs in all or part of the databases to synchronize and maintain all LSDBs.

  An SNP consists of a complete SNP (CSNP) and a partial SNP (PSNP). They are further divided into a Level-1 CSNP, Level-2 CSNP, Level-1 PSNP, and Level-2 P SNP.

  A CSNP contains the summary of all LSPs in an LSDB. This maintains LSDB synchronization between neighboring routers. On a broadcast network, the DIS periodically sends CSNPs. The default interval for sending CSNPs is 10 seconds. On a point-to-point link, CSNPs are sent only when the neighbor relationship is established for the first time.

  Figure 16-28 shows the CSNP packet format.

Figure 16-28 FLevel-1 or Level-2 CSNP packet

	No. of Octets
Intradomain Routeing Protocol Discriminator	1
Length Indicator	1
Version/Protocol ID Extension	1
ID Length	1
R R R PDU Type	1
Version	1
Reserved	1
Maximum Area Address	1
PDU Length	2
Source ID	ID Length+1
Start LSP ID	ID Length+2
End LSP ID	ID Length+2
Variable Length Fields	

The main fields of the CSNP packet are described as follows:

- Source ID: indicates the system ID of the router that sends the SNP.
- Start LSP ID: indicates the ID of the first LSP in the CSNP.
- End LSP ID: indicates the ID of the last LSP in the CSNP.

A PSNP lists only the sequence number of recently received LSPs. A PSNP can acknowledge multiple LSPs at one time. If an LSDB is not updated, the PSNP is also used to request a neighbor to send a new LSP.

Figure 16-29 shows the PSNP packet format.

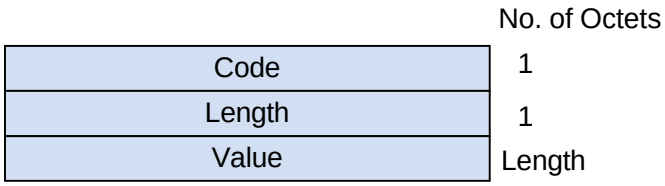
Figure 16-29 Level-1 or Level-2 PSNP packet

	No. of Octets
Intradomain Routeing Protocol Discriminator	1
Length Indicator	1
Version/Protocol ID Extension	1
ID Length	1
R R R PDU Type	1
Version	1
Reserved	1
Maximum Area Address	1
PDU Length	2
Source ID	ID Length+1
Variable Length Fields	

- CLV

  The variable length fields in a PDU are the multiple Code-Length-Values (CLVs). A CLV is also called the Type- Length-Value (TLV).Figure 16-30 shows the CLV format.

  Figure 16-30 CLV format



  CLVs vary with PDU types, as shown in Table 16-10.

Table 16-10 PDU types and CLV names

CLV Code	Name	Applied PDU Type
1	Area Addresses	IIH and LSP
2	IS Neighbors (LSP)	LSP
4	Partition Designated Level2 IS	L2 LSP
6	IS Neighbors (MAC Address)	LAN IIH
7	IS Neighbors (SNPA Address)	LAN IIH
8	Padding	IIH
9	LSP Entries	SNP
10	Authentication Information	IIH, LSP, and SNP
128	IP Internal Reachability Information	LSP
129	Protocols Supported	IIH and LSP
130	IP External Reachability Information	L2 LSP
131	Inter-Domain Routing Protocol Information	L2 LSP
132	IP Interface Address	IIH and LSP

  CLVs with codes 1 to 10 are defined in ISO 10589 (type 3 and 5) are not listed in the table. The other CLVs are defined in RFC 1195.

IS-IS Areas

- Two-Level structure

  To support large-scale routing networks, IS-IS adopts a two-level structure in a routing domain. A large domain can be divided into one or more areas. In general, Level-1 routers are located in an area, Level-2 routers are located among areas, and Level-1-2 routers are located between the Level-1 and Level-2 routers.

- Level-1 router

  A Level-1 router manages intra-area routing. It establishes neighbor relationships with only the Level-1 and Level-1-2 routers in the same area. It maintains a Level-1 LSDB. The LSDB contains routing information of the local area. A packet to a destination outside this area is forwarded to the nearest Level-1-2 router.

- Level-2 router

  A Level-2 router manages inter-area routing. It can establish neighbor relationships with Level-2 routers or Level-1-2 routers in other areas. It maintains a Level-2 LSDB. The LSDB contains inter-area routing information.

  All Level-2 routers form the backbone network of the routing domain. They are responsible for communications between areas. The Level-2 routers in the routing domain must be in succession to ensure the continuity of the backbone network. Only Level-2 routers can exchange data packets or routing information with routers outside the routing domain.

- Level-1-2 router

  A router that belongs to both a Level-1 area and a Level-2 area, is called a Level-1-2 router. It can establish Level-1 neighbor relationships with Level-1 routers and Level-1-2 routers in the same area. It can also establish Level-2 neighbor relationships with Level-2 routers and Level-1-2 routers in other areas. A Level-1 router must be connected to other areas through a Level-1-2 router.

  A Level-1-2 router maintains two LSDBs: Level-1 and Level-2. The Level-1 LSDB is used for intra-area routing and the Level-2 LSDB is used for inter-area routing.

NOTE

Level-1 routers in different areas cannot establish neighbor relationships. Level-2 routers can establish neighbor relationships with each other, regardless of the areas to which the Level-2 routers belong.

- Interface level

  A Level-1-2 router might need to establish only a Level-1 neighbor relationship with one remote end and only a Level-2 neighbor relationship with the other remote end. You can set the level of an interface to restrict the setup of adjacencies on the interface. For example, only a Level-1 adjacency can be established on a Level-1 interface and only a Level-2 adjacency can be established on a Level-2 interface.

Figure 16-31 shows a network that runs IS-IS. The network is similar to an OSPF network typology with multiple areas. The entire backbone area contains all routers in Area 1 and Level-1-2 routers in other areas.

Figure 16-31 IS-IS topology I

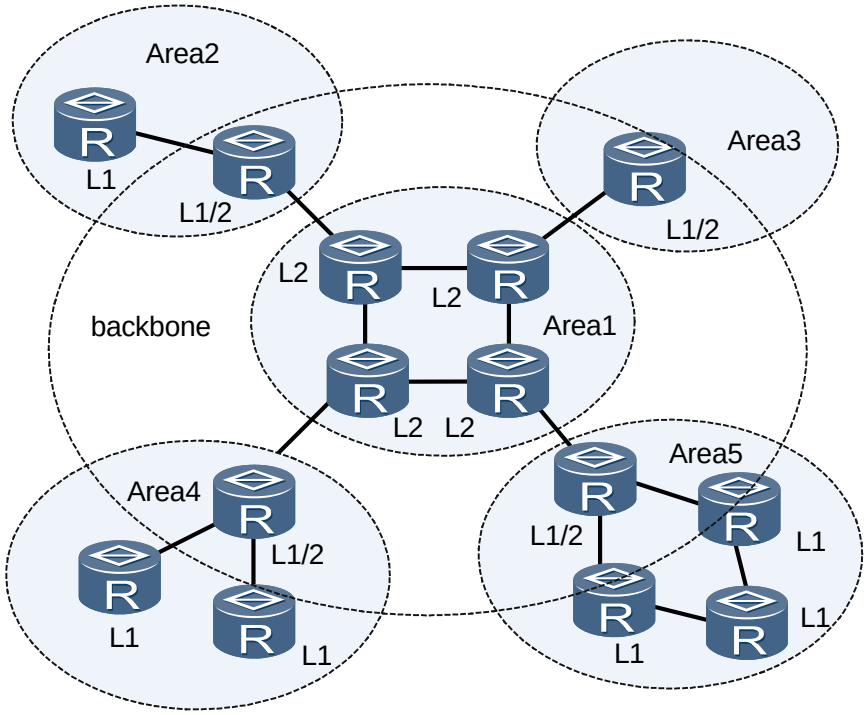
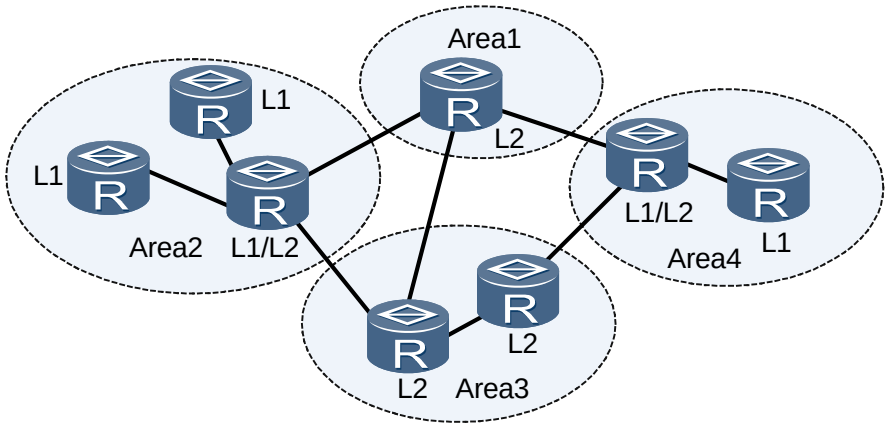


Figure 16-32 shows another type of IS-IS topology. All the successive Level-1-2 and Level-2 routers act as the backbone area of IS-IS. In this topology, Level-2 routers belong to different areas, and Level-1-2 routers also belong to different areas. No area is specifically defined as the backbone area.

Figure 16-32 IS-IS topology II



NOTE

The IS-IS backbone network does not refer to a specific area.

This networking scheme shows the difference between IS-IS and OSPF. For OSPF, inter-area routes are forwarded by the backbone area, and the SPF algorithm is used only in the same area. For IS-IS, both Level-1 and Level-2 routes are calculated through the SPF algorithm to generate the Shortest Path Tree (SPT).

IS-IS Network Types

IS-IS supports only two types of networks. In terms of physical links, IS-IS networks can be classified into the following link types:

- Broadcast: such as Ethernet and Token-Ring
- Point-to-point: such as PPP and HDLC

For a Non-Broadcast Multi-Access (NBMA) network such as the ATM, you should configure its sub-interfaces as P2P interfaces. IS-IS cannot run on Point to MultiPoint (P2MP) networks.

DIS and Pseudo Node

In a broadcast network, IS-IS needs to elect a Designated Intermediate System (DIS) from all the routers.

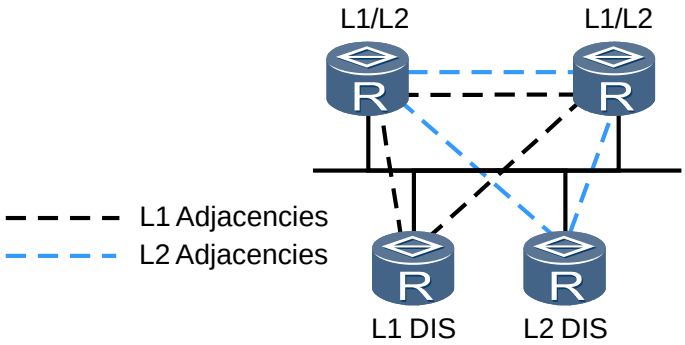
You can configure different priorities for DISs of different levels. The router with the highest priority is elected as the DIS. If there are multiple routers with the same highest priority in a broadcast network, the one with the highest MAC address is chosen. The DISs of different levels can be the same router or different routers.

Unlike DR election in OSPF, the DIS election in IS-IS has the following features:

- The router with the priority 0 also takes part in the DIS election.
- When a new router that meets the requirements to be a DIS joins the broadcast network, the router is selected as the new DIS, and the original pseudonode is deleted. This causes LSP flooding.

In an IS-IS broadcast network, the routers (including non-DIS routers) of the same level in a network segment set up adjacencies. This is different from an OSPF network. Figure 16-33 shows the networking for adjacencies.

Figure 16-33 DISs and adjacencies in an IS-IS broadcast network



A DIS is used to create and update pseudo nodes. It also generates LSPs of the pseudo nodes. The LSPs describe the available routers on the network.

The pseudo node is used to simulate the virtual node in the broadcast network and is not an actual router. In IS-IS, a pseudo node is identified by the system ID of the DIS and the 1-byte Circuit ID (its value is not 0).

With pseudo nodes, the network topology is simplified and LSPs are shortened. When the network changes, the number of generated LSPs is reduced. As a result, the SPF consumes fewer resources.

NOTE

In an IS-IS broadcast network, although all the routers set up adjacencies with each other, the LSDBs are synchronized by the DISs.

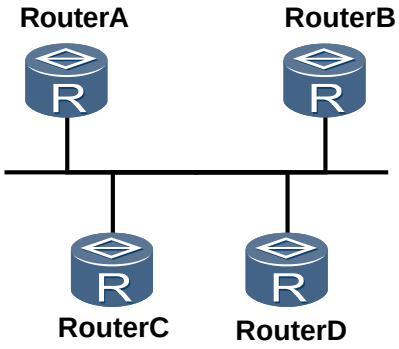
Establishment of IS-IS Neighbor Relationship

Two IS-IS routers need to establish a neighbor relationship before exchanging packets to implement routing. On different networks, the modes for establishing IS-IS neighbors are different.

- Establishment of a neighbor relationship on a broadcast link

  Figure 16-34 shows the neighbor relationship between Router A and Router B.

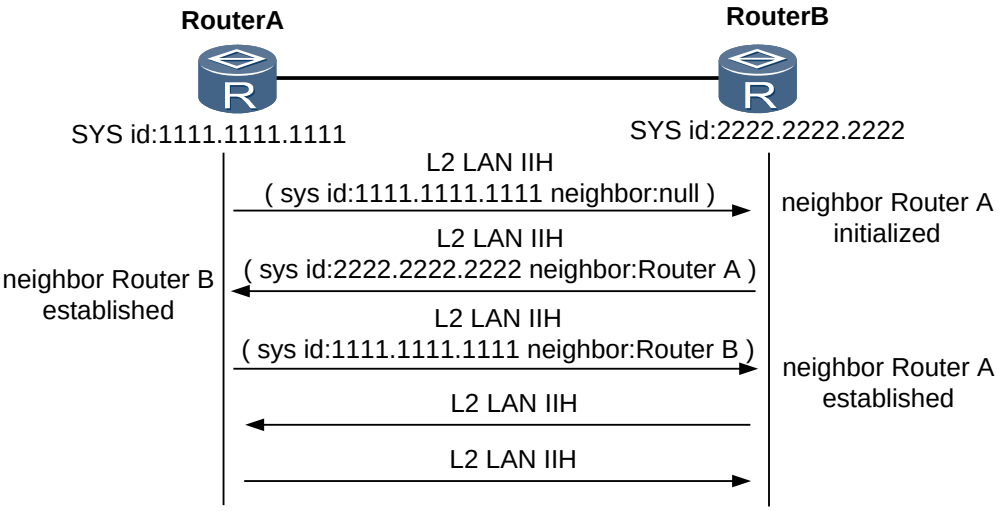
  Figure 16-34 Networking for a broadcast link



  Router A, Router B, Router C, and Router D are Level-2 routers. Router A is newly added to the broadcast network. The process of establishing the neighbor relationship between Router A and Router C or Router D is similar to that between Router A and Router B, and is not mentioned here.

  Figure 16-35 shows the process of establishing the neighbor relationship between Router A and Router B.

  Figure 16-35 Establishing neighbor relationship on a broadcast link



Router A broadcasts a Level-2 LAN IS-IS Hello PDU. After receiving the PDU, Router B sets its neighbor status with Router A to Initial. Then, Router B responds to Router A with a Level-2 LAN IIH packet indicating that Router A is a neighbor of Router B. On receiving the IIH packet, Router A sets its neighbor status with Router B to Up.

The network is a broadcast network, so a DIS needs to be elected. After the neighbor relationship is established, routers wait for two intervals before sending Hello packets to elect the DIS. The IIH packets exchanged by the routers contain the Priority field. The router with the highest priority is elected as the DIS. If the routers have the same priority, the router with the largest interface MAC address is elected as the DIS.

- Establishment of a neighbor relationship on a P2P link

  Unlike the establishment of a neighbor relationship on a broadcast link, the establishment of a neighbor relationship on a P2P link is classified into two modes: two-way mode and three-way mode.

  - Two-way mode

    Upon receiving an IS-IS Hello packet, a router unidirectionally sets up the neighbor relationship.

  - Three-way mode

    A neighbor relationship is established after IS-IS Hello PDUs are sent for three times, which is similar to the establishment of a neighbor relationship on a broadcast link.

NOTE
For details on three-way handshake mechanism of IS-IS, see 16.8.11 IS-IS Three-way Handshake chapters.

Basic rules for establishing an IS-IS neighbor relationship are as follows:

- Only neighboring routers of the same level can set up the neighbor relationship with each other.
- For Level-1 routers, their area IDs must be the same.
- Routers must be on the same network segment.

Network types of IS-IS interfaces on both ends of a link must be consistent. Otherwise, the neighbor relationship cannot be established. By simulating Ethernet interfaces as P2P interfaces, you can establish a neighbor relationship on a P2P link.

IS-IS runs on the data-link layer and was initially designed for CLNP. Therefore, the establishment of an IS-IS neighbor relationship is not related to IP addresses. In the implementation of a device, IS-IS runs only over IP. Therefore, IS-IS needs to check the IP address of its neighbor. If secondary IP addresses are assigned to the interfaces, the routers can still set up the IS-IS neighbor relationship, but only when either the primary IP addresses or secondary IP addresses are on the same network segment.

When IP address unnumbered is not configured, if the IP address of its neighbor and the address of the interface through which the router receives packets are not on the same network segment, the neighbor relationship cannot be set up, and IP unreachability is prevented. The neighbor relationship can be set up if you configure the router to not check the IP addresses contained in received Hello packets.

- For P2P interfaces, you can configure the interfaces not to check the IP addresses.
- For Ethernet interfaces, you must simulate Ethernet interfaces as P2P interfaces and then configure the interfaces to not check the IP addresses.

Process of Exchanging IS-IS LSPs

- LSP flooding

  LSP flooding is a mode in which a router sends an LSP to its neighbors and the neighbors send the received LSP to their respective neighbors except for the router that first sent the LSP. The LSP is flooded among the routers of the same level. Through flooding, each router of the same level receives the same LSP information and keeps a synchronized LSDB.

  Each LSP has a 4-byte sequence number. When a router is started, the sequence number of the first LSP sent by the router is 1. When a new LSP is generated, the sequence number of the LSP is equal to the sequence number of the previous LSP plus 1. The greater the sequence number, the newer the LSP.

- Causes of LSP generation

  All routers in the IS-IS routing domain can generate LSPs. The following events trigger the generation of a new LSP:

  - Neighbor is Up or Down.
  - Related interface is Up or Down.
  - Imported IP routes change.
  - Inter-area IP routes change.
  - Interface is assigned a new metric value.
  - Periodic updates occur.

- Processing of a new LSP received from a neighbor

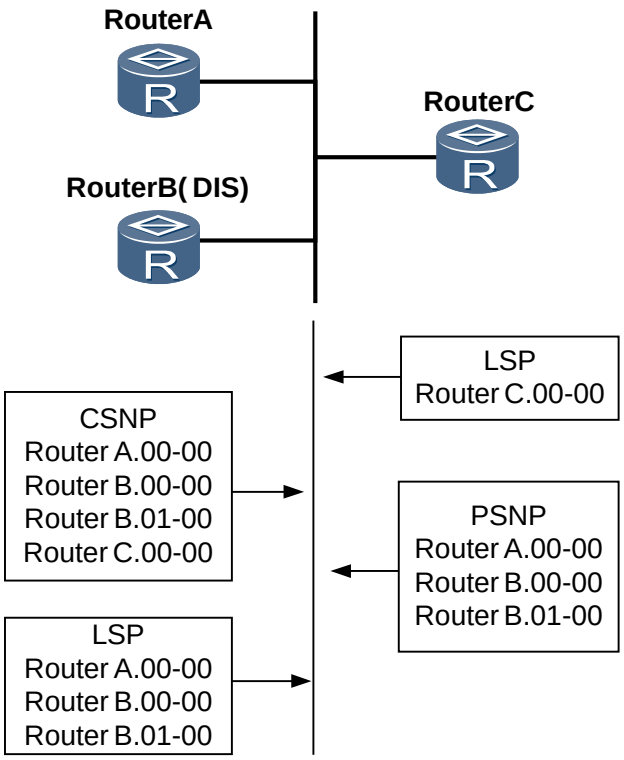
  a. The router installs the LSP to the LSDB and marks it for flooding.

  b. The router sends the LSP to all interfaces except the interface that initially received the LSP.

  c. The neighbors flood the LSP to their neighbors.

- Synchronizing LSDBs between a newly added router and DIS

Figure 16-36 Updating LSDBs on a broadcast link



- Newly added Router C sends Hello packets to establish neighbor relationships with the other routers in the broadcast domain. For details, see "Establishment of a neighbor relationship on a broadcast link."
- After setting up the neighbor relationships with other routers, Router C sends its LSP to the following multicast addresses after the LSP timer expires:

  Level-1: 01-80-C2-00-00-14

  Level-2: 01-80-C2-00-00-15

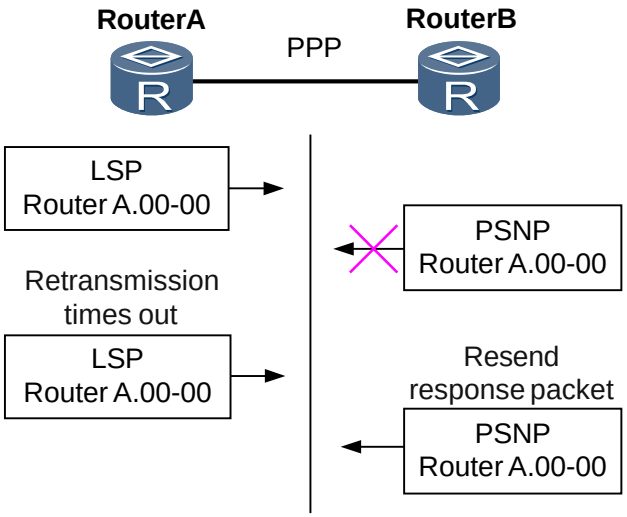
  Now, all neighbors on the network can receive the LSP.
- The DIS on the network segment adds the LSP received from Router C to its LSDB. After the CSNP timer expires, the DIS sends CSNPs to synchronize the LSDBs on the network. By default, CSNPs are sent at intervals of 10 seconds.
- After Router C receives the CSNPs from the DIS, Router C checks its LSDB and sends a PSNP to request the LSPs it does not have.
- After receiving the PSNP, the DIS sends the required LSPs to synchronize LSDBs.

- Process of updating the LSDB of the DIS
  - When the DIS receives an LSP, it searches the LSDB for the related records. If the DIS does not find the LSP in its LSDB, it adds the LSP to its LSDB and broadcasts the contents of the new LSDB.
  - If the sequence number of the received LSP is greater than the sequence number of the corresponding LSP, the DIS replaces the LSP with the received LSP in the LSDB, and broadcasts the contents of the new LSDB.
  - If the sequence number of the received LSP is smaller than the sequence number of the corresponding LSP, the DIS sends the LSP in the LSDB to the inbound interface.

- If the sequence number of the received LSP is equal to the sequence number of the corresponding LSP, the DIS compares the Remaining Lifetime of the two LSPs. If the received LSP has a smaller Remaining Lifetime than that of the corresponding LSP, the DIS replaces the LSP with the received LSP, and broadcasts the contents of the new LSDB. If the received LSP has a greater Remaining Lifetime than that of the corresponding LSP, the DIS sends the LSP to the inbound interface.
- If both the sequence number and the Remaining Lifetime of the received LSP and the corresponding LSP in the LSDB are the same, the DIS compares the checksum of the two LSPs. If the received LSP has a greater checksum than that of the corresponding LSP in the LSDB, the DIS replaces the LSP in the LSDB with the received LSP, and advertises the contents of the new LSDB. If the received LSP has a smaller checksum than that of the corresponding LSP in the LSDB, the DIS sends the LSP in the LSDB to the inbound interface.
- If the sequence number, Remaining Lifetime, and checksum of the received LSP and that of the corresponding LSP are the same, the LSP is not forwarded.

- Synchronizing the LSDB on a P2P link

Figure 16-37 Updating the LSDB on a P2P link



a. When the neighbor relationship is set up for the first time, a router sends a CSNP to its neighbor. If the LSDB of the neighbor and the CSNP are not synchronized, the neighbor sends PSNP requests for a required LSP.
b. The router sends the required LSP to the neighbor and starts the LSP retransmission timer. The router then waits for a PSNP from the neighbor as an acknowledgement receiving the LSP.
c. If the router does not receive the PSNP from the neighbor after the LSP retransmission timer expires, it resends the LSP.

NOTE

A PSNP on a P2P link is used as follows:

- An Ack packet to acknowledge the received LSP.
- A request packet to acquire LSPs.

- Updating the LSDB of the P2P
  - If the sequence number of the received LSP is greater than the sequence number of the corresponding LSP in the LSDB, the router adds the LSP to its LSDB. The

router then sends a PSNP to acknowledge the received LSP. At last, the router sends the LSP to all its neighbors except the neighbor that sends the LSP.

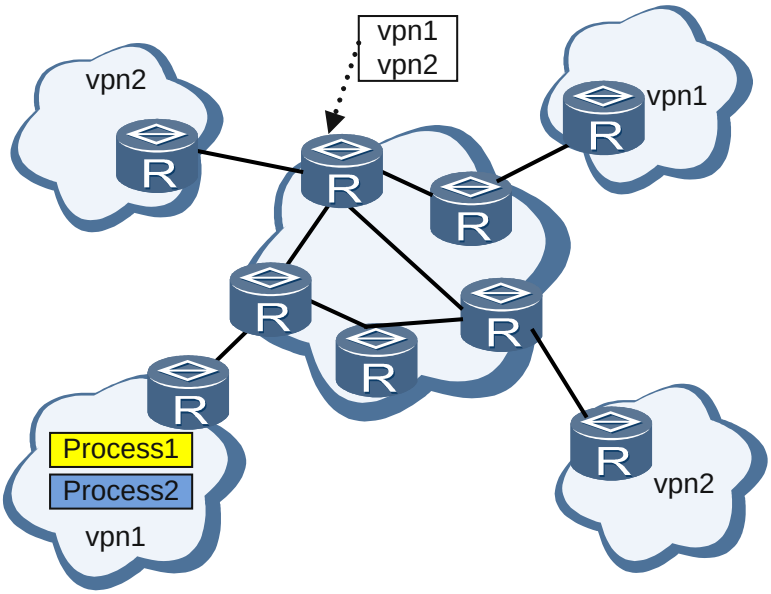
- If the sequence number of the received LSP is smaller than the sequence number of the corresponding LSP, the router directly sends its LSP to the neighbor and waits for a PSNP from the neighbor as the acknowledgement.
- If the sequence number of the received LSP is the same as the sequence number of the corresponding LSP in the LSDB, the router compares the Remaining Lifetime of the two LSPs. If the received LSP has a smaller Remaining Lifetime than that of the corresponding LSP, the router adds the LSP to its LSDB. The router then sends a PSNP to acknowledge the received LSP. If the received LSP has a greater Remaining Lifetime than that of the corresponding LSP in the LSDB, the router directly sends its LSP to the neighbor and waits for a PSNP from the neighbor.
- If both the sequence number and the Remaining Lifetime of the received LSP and the corresponding LSP in the LSDB are the same, the router compares the checksum of the two LSPs. If the received LSP has a greater checksum than that of the corresponding LSP, the router adds the LSP to its LSDB. The router then sends a PSNP to acknowledge the received LSP. If the received LSP has a smaller checksum than that of the corresponding LSP, the router directly sends its LSP to the neighbor and waits for a PSNP from the neighbor. At last, the router sends the LSP to all its neighbors except the neighbor that sends the LSP.
- If both the sequence number, Remaining Lifetime, and checksum of the received LSP and the corresponding LSP are the same, the LSP is not forwarded.

16.8.3 IS-IS Multi-instance and Multi-process

For the routers that support the VPN, you can associate each IS-IS process with a specific VPN instance. Therefore, you can configure multiple IS-IS processes to be associated with multiple VPN instances at the same time.

- IS-IS multi-instance indicates that you can configure multiple IS-IS instances on the same router.
- IS-IS multi-process indicates that you can create multiple IS-IS processes in a VPN or a public network. As shown in Figure 16-38.
  - The multi-process feature allows a set of interfaces to be associated with a specific IS-IS process. This ensures that the specific IS-IS process performs all the protocol operations only on the set of interfaces. Therefore, multiple IS-IS processes can work on a single router and each process is responsible for a unique set of interfaces.
  - IS-IS multi-processes share an RM routing table. IS-IS multi-instances use the RM routing tables of VPNs. Each VPN has its own RM routing table.
  - When an IS-IS process is created, it can be associated with a VPN instance. Then, the IS-IS process belongs to the VPN and processes events only in the VPN. The IS-IS process is deleted when the associated VPN is deleted.

Figure 16-38 Networking diagram for IS-IS multi-instance and multi-process



For easy management and effective control, IS-IS supports multi-process and multi-instance features.

In the scenario where IS-IS is applied to users on private networks, after a VPN is created, interfaces bound to the VPN and routes in the VPN are isolated from other VPNs and public network data. In this case, you can adopt IS-IS multi-instance to deploy IS-IS in the VPN.

For the routers that support the VPN, each IS-IS process is associated with a specific VPN instance. All the interfaces attached to an IS-IS process, therefore, should be associated with the VPN instance that this IS-IS process is associated to.

At present, the VPN instance is maintained by the VPN module. Therefore, IS-IS multi-instance is implemented by associating an IS-IS process with a VPN instance when the IS-IS process is created.

When configuring IS-IS multi-instance and multi-process, note the following:

- When creating IS-IS multi-instances, associate an IS-IS process with a VPN instance when the IS-IS process is created. If an IS-IS process is not associated with a VPN instance when the IS-IS process is created, the association cannot be configured later.
- An IS-IS process that is already associated with a VPN instance cannot be associated with another VPN instance.
- Multiple IS-IS processes can be associated with one VPN instance.
- The interfaces where IS-IS multi-instance needs to be enabled must be associated with the same VPN instance as IS-IS.
- The IS-IS process associated with a VPN instance belongs to the VPN. Therefore, the IS-IS process is deleted when the VPN is deleted.
- Routes from different VPNs cannot be imported to each other.

16.8.4 IS-IS Route Leaking

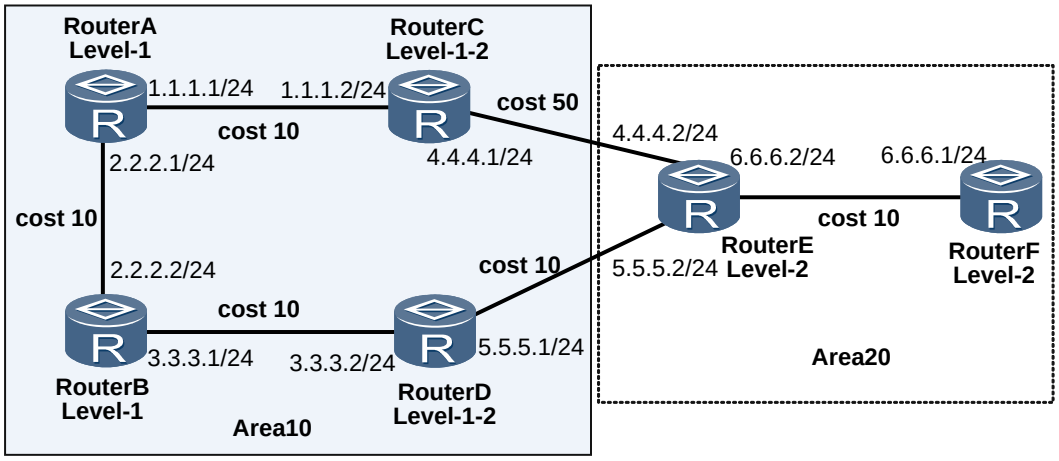
With the route leaking function, Level-1-2 IS-IS advertises to the specified Level-1 areas the known routing information about other Level-1 and Level-2 areas.

Normally, intra-area routes are managed by Level-1 routers. All Level-2 and Level-1-2 routers form a successive backbone area. The Level-1 area can be connected to only the backbone area. The Level-1 areas cannot be connected to each other.

Routing information in a Level-1 area is advertised to a Level-2 area through a Level-1-2 router. That is, the Level-1-2 router encapsulates the learned Level-1 routing information into a Level-2 LSP and floods the Level-2 LSP to other Level-2 and Level-1-2 routers. To reduce the size of routing tables, Level-2 routers, by default, do not advertise the learned routing information of Level-1 areas and that of the backbone area to Level-1 areas. Because Level-1 routers cannot access the routing information outside the area. The Level-1 routers cannot select the optimal route to a destination.

IS-IS route leaking can solve this problem. By configuring Access Control Lists (ACLs) and route-policies and marking routes with tags on Level-1-2 routers, you can select eligible routes. Then, a Level-1-2 router can advertise routing information of other Level-1 areas and backbone area to its Level-1 area.

Figure 16-39 Networking for route leaking



- In the figure, Router A, Router B, Router C, and Router D belong to Area 10. Router A and Router B are Level-1 routers; Router C and Router D are Level-1-2 routers.
- Router E and Router F are Level-2 routers and belong to Area 20.

If Router A sends a packet to Router F, the selected optimal route should be Router A -> Router B -> Router D -> Router E -> Router F, which has a cost of 40. However, if you check the route on Router A to view the path of packets sent to Router F, you can find that the selected route is Router A -> Router C -> Router E -> Router F, of which the cost is 70. This route is not the optimal route from Router A to Router F.

Because Router A does not detect the routes outside the local area, Router A sends the packets to other network segments through the default route generated by the nearest Level-1-2 router.

To avoid this scenario, you can enable route leaking on the Level-1-2 routers Router C and Router D. When you check the route, you find that the selected route is Router A -> Router B -> Router D -> Router E -> Router F, which is the optimal route from Router A to Router F.

16.8.5 IS-IS Fast Convergence

IS-IS fast convergence is an extended feature of IS-IS that is implemented to speed up the convergence of routes. Fast convergence includes the following:

- Incremental SPF (I-SPF): recalculates only the routes of the changed nodes rather than all the nodes when the network topology changes. This speeds up the calculation of routes.
- Partial Route Calculation (PRC): calculates only the changed routes when the routes on the network change.
- LSP fast flooding: speeds up the flooding of LSPs.
- Intelligent timer: is applicable to LSP generation and SPF calculation.

  The first timeout period of the timer is fixed. If an event that triggers the timer happens while the timer is set and unexpired, intelligent timer increases the interval it sets for next time.

I-SPF

In ISO 10589, the Dijkstra algorithm was adopted to calculate routes. When a node changes on the network, this algorithm is used to recalculate all routes. The calculation takes a long time and consumes too many CPU resources, which affects the convergence speed.

I-SPF improves this algorithm. Except for the first time, only changed nodes instead of all nodes are involved in calculation. The shortest path tree (SPT) generated is the same as that generated by the previous algorithm. This decreases CPU usage and speeds up network convergence.

PRC

Similar to I-SPF, PRC calculates only the changed routes, but it does not calculate the shortest path. It updates routes based on the SPT calculated by I-SPF.

In route calculation, a leaf represents a route, and a node represents a router. If the SPT changes after I-SPF calculation, PRC processes all the leaves only on the changed node. If the SPT remains unchanged, PRC processes only the changed leaves.

For example, if IS-IS is enabled on an interface of a node, the SPT calculated by I-SPF remains unchanged. PRC updates only the routes of this interface, consuming less CPU resources.

PRC working with I-SPF further improves the convergence performance of the network. It is an improvement of the original SPF algorithm.

NOTE
In the implementation of a device, only I-SPF and PRC are used to calculate IS-IS routes.

LSP Fast Flooding

When IS-IS receives new LSPs from other routers, it updates the LSPs in the LSDB and periodically floods out the updated LSPs based on a timer.

LSP fast flooding improves on the PRC mode. When the device configured with this feature receives one or more new LSPs, before it calculates routes, it floods out the LSPs whose amount is smaller than the specified number. Network convergence speed is significantly improved.

Intelligent Timer

Although the route calculation algorithm is improved, the long interval for triggering route calculation affects the convergence speed. Frequent network changes also consume too many CPU resources. The SPF intelligent timer addresses both of these problems.

In general, an IS-IS network is stable under normal conditions. The probability of the occurrence of many network changes is very minimal, and IS-IS does not calculate routes frequently. The period for triggering the route calculation is very short (milliseconds). If the topology of the network changes very often, the intelligent timer increases the interval for the calculation times to avoid too much CPU consumption. The original mechanism uses a timer with uniform intervals, which makes fast convergence and low CPU consumption impossible to achieve.

The LSP generation intelligent timer is similar to the SPF intelligent timer. When the LSP generation intelligent timer expires, the system generates a new LSP based on the current topology. The LSP generation timer is designed as an intelligent timer to respond to emergencies (such as the interface is Up or Down) quickly and speed up the network convergence.

16.8.6 Priority-based IS-IS Convergence

Priority-based IS-IS convergence ensures that specific routes converge first in the case of a great number of routes. Different routes can be set with different convergence priorities.

Priority-based IS-IS convergence enables specific routes (such as routes that match the specified IP prefix) to converge first. You can assign a high convergence priority to routes for key services so that these routes converge quickly. This decreases impact on key services and improves network reliability.

16.8.7 IS-IS LSP Fragment Extension

When the LSPs to be advertised by IS-IS contain much information, they are advertised in multiple LSP fragments of the same system. The IS-IS LSP fragment extension attribute allows an IS-IS router to generate more LSP fragments and carry more IS-IS information.

As defined in RFC 3786, virtual system IDs can be configured and virtual LSPs that carry routing information can be generated for IS-IS.

Terms

- Originating system: is a router that runs the IS-IS protocol. A single IS-IS process can advertise its LSPs as multiple "virtual" routers, and the originating system represents the "real" IS-IS process.
- Normal system ID: is the system ID of the originating system.
- Additional system ID: assigned by network administrators, is used to generate additional or extended LSP fragments. Up to 256 additional or extended LSP fragments can be generated. Like the normal system ID, the additional system ID must be unique in the routing domain.

  The additional system ID, assigned by network administrators, is used to generate additional or extended LSP fragments. Up to 256 additional or extended LSP fragments can be generated. Like the normal system ID, the additional system ID must be unique in the routing domain.
- Virtual system: identified by an additional system ID, is used to generate extended LSP fragments. These fragments carry the additional system IDs in their LSP IDs.

Principle

IS-IS LSP fragments are identified by the LSP Number field in their LSP IDs. The LSP Number field is 1 byte. An IS-IS process can generate a maximum of 256 fragments that carry a limited number of routes (when the fragment length is 1497 bytes, a maximum of 30,000 routes can be carried). With fragment extension, more information can be carried.

With additional system IDs (up to 50 virtual systems), an IS-IS process can generate a maximum of 13056 LSP fragments.

When a virtual system and fragment extension are configured, an IS-IS router adds the contents that cannot be contained in the LSPs advertised by the originating system to the LSPs of the virtual system. The router notifies other routers of the relationship between the virtual system and itself through a special TLV.

IS Alias ID TLV

A special TLV, IS Alias ID TLV, is defined in RFC 3786.

Table 16-11 IS Alias ID TLV

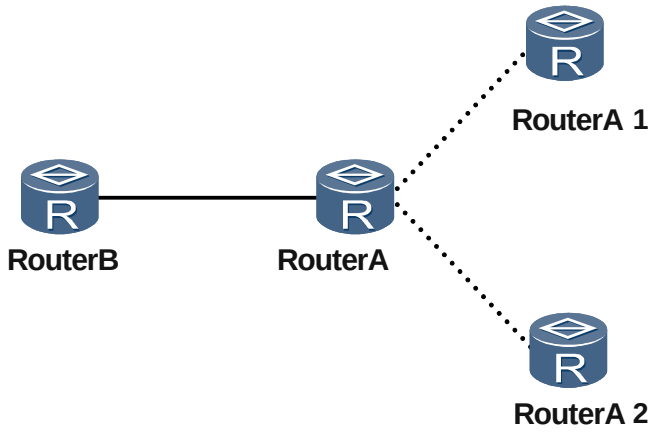
Field	Length	Description
Type	1 byte	Indicates the TLV type. If the value is 24, it indicates the IS Alias ID TLV.
Length	1 byte	Indicates the length of Value in the TLV.
System ID	6 bytes	System ID
Pseudonode number	1 byte	pseudonode number
sub-TLVs length	1 byte	sub-TLVs length
sub-TLVs	0 to 247 bytes	sub-TLVs

Regardless of the operation mode, the originating system and virtual system send the LSPs with fragment number 0 carrying the IS Alias ID TLV to indicate the originating system.

Operation Modes

Figure 1 shows the networking for the LSP fragment extension feature, which can be run in two different modes.

Figure 16-40 LSP fragment extension



- The IS-IS router can run the LSP fragment extension feature in the following modes:

  Mode-1: is used when some routers on the network do not support the LSP fragment extension.

  In this mode, virtual systems participate in the SPF calculation. The originating system advertises LSPs that contain information about the links to each virtual system. Similarly, each virtual system advertises LSPs that contain information about the links to the originating system. This allows the virtual systems to appear to be like the actual routers connected to the originating system on the network.

  Mode-1 is a transitional mode for earlier IS-IS versions that do not support fragment extension. In the earlier versions, IS-IS cannot identify the Alias ID TLV. The LSP sent by a virtual system appears to be like a common IS-IS LSP.

  The LSP sent by a virtual system contains the same area address and overload bit as that in the common LSP. If the LSPs sent by a virtual system contain TLVs specified in other features, they must be the same as those in common LSPs.

  The virtual system carries neighbor information that specifies that the neighbor is the originating system, with the metric being the maximum value minus 1. The originating system carries neighbor information that specifies that the neighbor is the virtual system, with the metric of 0. This ensures that the virtual system is the downstream node of the originating system when other routers calculate routes.

  As shown in Figure 16-40, Router B does not support the LSP fragment extension; Router A is set to support the LSP fragment extension in mode-1; Router A1 and Router A2 are virtual systems of Router A. In this example, Router A1 and Router A2 send LSPs carrying routing information of Router A. After receiving LSPs from Router A, Router A1, and Router A2, Router B detects that there are three individual routers at the peer end and calculates routes normally. Because the cost of the route from Router A to Router A1 and the cost of the route from Router A to Router A2 are both 0, the cost of the route from Router B to Router A is equal to the cost of the route from Router B to Router A1.

- Mode-2: is used when all routers on the network support LSP fragment extension.

  In this mode, virtual systems do not participate in the SPF calculation. All routers on the network detects that the LSPs generated by the virtual systems actually belong to the originating system.

  Working in mode-2, IS-IS identifies IS Alias ID TLV, which is used to calculate the SPT and routes.

As shown in Figure 16-40, Router B supports the LSP fragment extension; Router A is set to support the LSP fragment extension in mode-2; andRouter A1 and Router A2 send LSPs carrying routing information of Router A. When receiving LSPs from Router A1 and Router A2, Router B obtains IS Alias ID TLV and detects that the originating system of Router A1 and Router A2 is Router A. Router B detects that information advertised by Router A1 and Router A2 belongs to Router A.

Whether LSP fragment extension is set to mode-1 or mode-2, LSPs in both modes can be resolved. If LSP fragment extension is not supported, only LSPs in mode-1 can be resolved.

Table 16-12 Comparison between LSP fragment extension mode-1 and mode-2

LSP Content\Mode	Mode-1	Mode-2
IS Alias ID	Yes	Yes
area	Yes	No
overload bit	Yes	Yes
IS NBR/IS EXTENDED NBR	Yes	No
Routing	Yes	Yes
ATT bits	must be 0	must be 0
P bit	must be 0	must be 0

Process

After LSP fragment extension is configured, if information is lost because LSPs are of full lengths, the system prompts that the IS-IS router should be restarted. After the router is restarted, the originating system loads as much routing information as possible. The remaining information is added to the LSPs of the virtual systems for transmission.

Application Environment

NOTE

If there are devices of other manufacturers on the network, LSP fragment extension must be set to mode-1. Otherwise, devices of other manufacturers cannot identify the LSPs.

Configure the LSP fragment extension and virtual systems before you set up IS-IS neighbors or import routes. Then you must restart the IS-IS router for the configurations to take effect. If you set up IS-IS neighbors or import routes first, it can cause IS-IS to carry more information than cannot be loaded through 256 fragments

16.8.8 IS-IS Administrative Tag

Administrative tags carry information about IP address prefixes and control advertisement of IP prefixes in the IS-IS domain. They are used to control the importing of routes of different levels and areas, and to control different routing protocols and IS-IS multi-instances running on the same router.

The value of an administrative tag is associated with certain attributes. If the cost-style is wide, wide-compatible or compatible, when IS-IS advertises an IP address prefix with these

attributes, IS-IS adds the administrative tag to the TLV in the prefix. The tag is flooded along with the prefix throughout the routing domain.

16.8.9 Dynamic Hostname Exchange Mechanism

The dynamic hostname exchange mechanism provides a mapping from the hostname to system ID for IS-IS routers.

On an IS-IS router without hostname exchange, information about IS-IS neighbors and LSDBs is represented by a system ID with 12 hexadecimal numbers, for example, aaaa.eeee.1234. This representation is complicated and not easy to use.

To easily maintain and manage IS-IS networks easily, the dynamic hostname exchange mechanism was introduced.

Dynamic hostname information is advertised in the form of a dynamic hostname TLV (type 137) in LSPs. The dynamic hostname exchange mechanism also provides a service to associate a host name with the Designated IS (DIS) on a broadcast network. Then, this mechanism advertises this association through LSPs in the form of a dynamic hostname TLV.

In the implementation of MA5600T/MA5603T/MA5608T, routers with IS-IS dynamic hostname mapping enabled add the Dynamic Hostname TLV (TLV type 137) that records the local host name to the LSPs they generate before sending out the LSPs.

Dynamic Hostname TLV (TLV type 137) includes the following fields:

- Type: indicates the dynamic hostname exchange mechanism.
- Length: indicates the total length of the value field.
- Value: indicates a character string of 1 to 255 characters.

The Dynamic Hostname TLV is optional and can be inserted anywhere in an LSP. The hostname value cannot be null. A router determines whether to carry the TLV in LSPs it sends. The router that receives the LSPs determines whether to ignore the TLV or obtain the TLV for its mapping table.

Implementation

- Matching rules

  The dynamic hostname mechanism abides by the longest match rule. First, System ID+NSEL is first compared. If that does not match, the system ID is then compared.
- Transmission of dynamic hostname

  The dynamic hostname can be carried by the original LSP only.
- Transmission of DIS dynamic hostname

  The DIS dynamic hostname is transmitted through the LSPs generated by the DIS.
- Priority of dynamic hostname

  The dynamic hostname takes precedence over the static hostname. When both a dynamic hostname and a static hostname are configured, the dynamic hostname replaces the static hostname.
- Configuration and resolution of dynamic hostname

  The dynamic hostname can be up to 64 bytes in length and a maximum of 255-byte contents can be resolved.

Application Environment

In maintenance and management, the hostname is easier to identify and retain than the system ID. After this function is configured, the hostname instead of the system ID is displayed when you view information about IS-IS on the router.

The hostname exchange mechanism implemented on the MA5600T/MA5603T/MA5608T includes dynamic hostname mapping and static hostname mapping. The system ID is replaced by the hostname in the following cases:

- When an IS-IS neighbor is displayed, the system ID of the IS-IS neighbor is replaced by the dynamic hostname. If the IS-IS neighbor is the DIS, then the system ID of the DIS is replaced by the dynamic hostname of the neighbor.
- When an LSP in the IS-IS LSDB is displayed, the system ID in the LSP ID is replaced by the dynamic hostname of the router that advertises the LSP.
- When details about the IS-IS LSDB are displayed, the Host Name field is included for the LSP generated by the router where dynamic hostname exchange is enabled; the system ID is replaced by the dynamic hostname of the IS-IS neighbor.

16.8.10 IS-IS HA

IS-IS HA includes hot standby, data backup, command line backup, batch backup, and real-time backup.

IS-IS backs up data from the Active Main Board (AMB) to the Standby Main Board (SMB). Whenever the AMB fails, the SMB becomes active and takes over the AMB. IS-IS, therefore, can keep working normally.

Basic Concepts

- Data backup

  It indicates backup of data of processes and interfaces.
- Command line backup

  If the AMB processes successfully, it sends the command lines to the SMB for processing. If the AMB fails to process, it records in the log that the command lines fail to take effect and does not send the command lines to the SMB for processing. If the SMB fails to process, the failure is recorded in the log.

Hot Standby

The IS-IS Hot Standby (HSB) feature is supported on the devices with a distributed structure.

In the running process of IS-IS HSB, IS-IS configurations on the AMB and those on the SMB are consistent. When the AMB/SMB switchover occurs, IS-IS on the new AMB performs GR. The new AMB resends a request for setting up the neighbor relationship to neighbors to synchronize LSDBs. Traffic, therefore, is not affected.

Batch Backup

- Backing up data in batches

  When the SMB is installed, all data of the AMB is backed up to the SMB. No configuration can be changed during batch backup.
- Backing up command lines in batches

When the SMB is installed, all configurations of the AMB are backed up to the SMB at a time. No configuration can be changed during batch backup.

Real-time Backup

- Real-time backup of data

  It indicates real-time backup of changed data of processes and interfaces to the SMB.
- Real-time backup of command lines

  It indicates that command lines that are run successfully on the AMB are backed up to the SMB.

16.8.11 IS-IS Three-way Handshake

IS-IS introduces the three-way handshake mechanism on P2P links to ensure a reliable data link layer.

According to ISO 10589, the two-way handshake mechanism of IS-IS uses Hello packets to set up P2P adjacencies between neighbors. When the router receives a Hello packet from its peer, it detects the status of the peer as Up and sets an adjacency with the peer. This mechanism has distinct disadvantages.

When two or more links exist between two routers, an adjacency can still be set up when one link is Down and the other is Up in the same direction. The router does not detect any faults on the link that is in the Down state, and still attempt to forward packets through this link.

The three-way handshake mechanism solves these problems on P2P links. In three-way handshake mode, the router detects the neighbor as Up only after confirming that the neighbor received the packet that it sent, and then sets up an adjacency with the neighbor.

In addition, the three-way handshake mechanism uses the 32-bit Extended Local Circuit ID field. This extends the original 8-bit Extended Local Circuit ID field and P2P links increase to more than 255 in number.

NOTE

The 3-way handshake mechanism of IS-IS is implemented on P2P links by default, as defined in RFC 3373.

16.8.12 IS-IS GR

IS-IS Graceful Restart (GR) is one of the high availability (HA) technologies. IS-IS GR extends IS-IS to support the GR capability to implement non-stop forwarding. RFC 3847 defines the IS-IS GR standard.

Because IS-IS is a link state routing protocol, all routers in an area must maintain the same network topologies, that is, the same LSDBs.

After the master/slave switchover, no neighbor information is stored on the restarted router. The first Hello packets sent by the router after restart do not contain the neighbor list. After receiving the Hello packets, the neighbor checks the two-way neighbor relationship and detects that it is not in the neighbor list of the Hello packets sent by the router. The neighbor relationship is interrupted.

The neighbor then generates new LSPs and floods the topology changes to all other routers in the area. Routers in the area calculate routes based on the new LSDBs, which leads to route interruption or routing loops.

Because no LSDB was stored on the restarted router, the router needs to synchronize its LSDB with those of the neighbors.

When restarting IS-IS without GR mode, IS-IS neighbor relationships are reset and LSPs are regenerated and flooded. This triggers the SPF calculation in the entire area, which causes route flapping and forwarding interruption in the area.

The IETF defined the GR standard, RFC 3847, for IS-IS. The restart of the protocol is processed for both the reserved FIB tables and unreserved FIB tables. Therefore, the route flapping and interruption of the traffic forwarding caused by the restart can be avoided.

When a router fails, neighbors at the routing protocol layer detect that their neighbor relationships are Down and then become Up again after a period. This is the *flapping* of neighbor relationships. The flapping of neighbor relationships causes route flapping, which leads to black hole routes on the restarted router or causes data services from the neighbors to be looped on the restarted router. This decreases the reliability on the network. GR was introduced to address route flapping.

Basic Concepts of IS-IS GR

IS-IS GR involves two roles, namely, GR restarter and GR helper.

- GR restarter: is the router that restarts in GR mode.
- GR- helper: is another GR router that helps the restarter to complete the GR process. The GR restarter must have the capability of the GR helper.

NOTE
By default, the device supports the GR helper.

To implement GR, IS-IS introduces the restart Type-Length-Value (TLV), T1 timer, T2 timer, and T3 timer.

Restart TLV

The restart TLV is an extended part of an IS-to-IS Hello (IIH) PDU. All IIH packets of the router that supports IS-IS GR contain the restart TLV. The restart TLV carries the parameters for the protocol restart. Figure 16-41 shows the format of the restart TLV.

Figure 16-41 Restart TLV

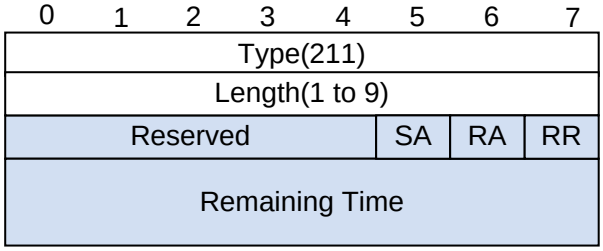


Table 16-13 describes the fields of the restart TLV.

Table 16-13 Restart TLV fields

Field	Length	Description

Field	Length	Description
Type	1 byte	TLV type. Type value 211 indicates the restart TLV.
Length	1 byte	Length of value in the TLV.
RR	1 bit	Restart request bit. A router sends an RR packet to notify the neighbors of its restarting or starting and to require the neighbors to retain the current IS-IS adjacency and return CSNPs.
RA	1 bit	Restart acknowledgement bit. A router sends an RA packet to respond to the RR packet.
SA	1 bit	Suppress adjacency advertisement bit. The starting router uses an SA packet to require its neighbors to suppress the broadcast of their neighbor relationships to prevent routing loops.
Remaining Time	2 bytes	Time during which the neighbor does not reset the adjacency. The length of the field is 2 bytes. The time is measured in seconds. When RA is reset, the value is mandatory.

Timers

Three timers are introduced to enhance IS-IS GR: T1, T2, and T3.

- T1

  Any interface enabled with IS-IS GR maintains a T1 timer. On a Level-1-2 router, broadcast interfaces maintain a T1 timer for Level-1 and Level-2 neighbor relationships.

  If the GR restarter has already sent an IIH packet with RR being set but does not receive any IIH packet that carries the restart TLV and the RA set from the GR helper even after the T1 timer expires, the GR restarter resets the T1 timer and continues to send the restart TLV.

  If the ACK packet is received or the T1 timer expires three times, the T1 timer is deleted. The default value of a T1 timer is 3 seconds.

- T2

  Level-1 and Level-2 LSDBs maintain separate T2 timers.

  T2 is the maximum time that the system waits for the synchronization of various LSDBs. T2 is generally 60 seconds.

- T3

  The entire system maintains a T3 timer.

  T3 can be considered the maximum time for GR to complete.

  If the T3 timer expires, GR fails.

  The initial value of the T3 timer is 65535 seconds. After the IIH packets that carry the RA are received from neighbors, the value of the T3 timer becomes the smallest value of the Remaining Time field among the Remaining Time fields of the IIH packets.

  The T3 timer only applies when devices are restarted.

Session Mechanism of IS-IS GR

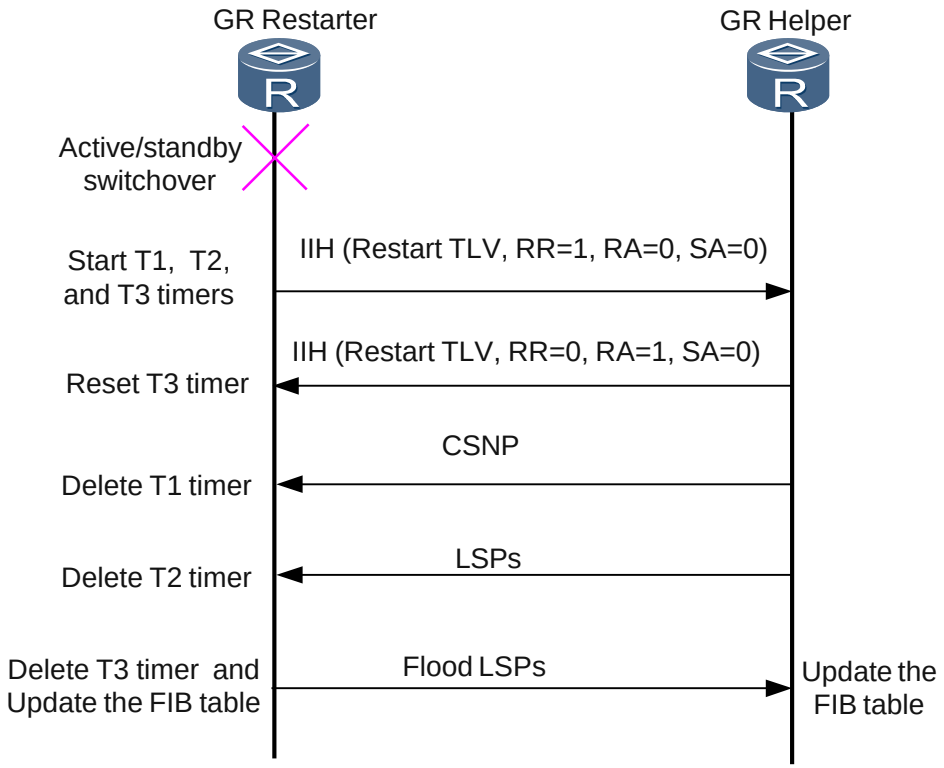
For differentiation, GR triggered by the master/slave switchover or the restart of an IS-IS process is referred to as *restarting*. In restarting, the FIB table remains unchanged. GR triggered by router restart is referred to as *starting*. In starting, the FIB table is updated.

The following describes the process of IS-IS GR in restarting and starting modes:

IS-IS Restarting

Figure 16-42 shows the process of IS-IS restarting.

Figure 16-42 IS-IS restarting



1. After performing the protocol restart, the GR restarter performs the following actions:
   - Starts T1, T2, and T3 timers.
   - Sends IIH packets that contain the restart TLV from all interfaces. In such a packet, RR is set to 1, and RA and SA are set to 0.
2. After receiving an IIH packet, the GR helper performs the following actions:
   - Maintains the neighbor relationship and refreshes the current Holdtime.
   - Replies with an IIH packet containing the restart TLV. In the packet, RR is set to 0; RA is set to 1, and the value of the Remaining Time field indicates the period from the current moment to the timeout of the Holdtime.
   - Sends CSNPs and all LSPs to the GR restarter.

NOTE

- On a P2P link, a neighbor must send CSNPs.

- On a LAN link, only the neighbor of the DIS sends CSNPs. If the DIS is restarted, a temporary DIS is elected from the other routers on the LAN.

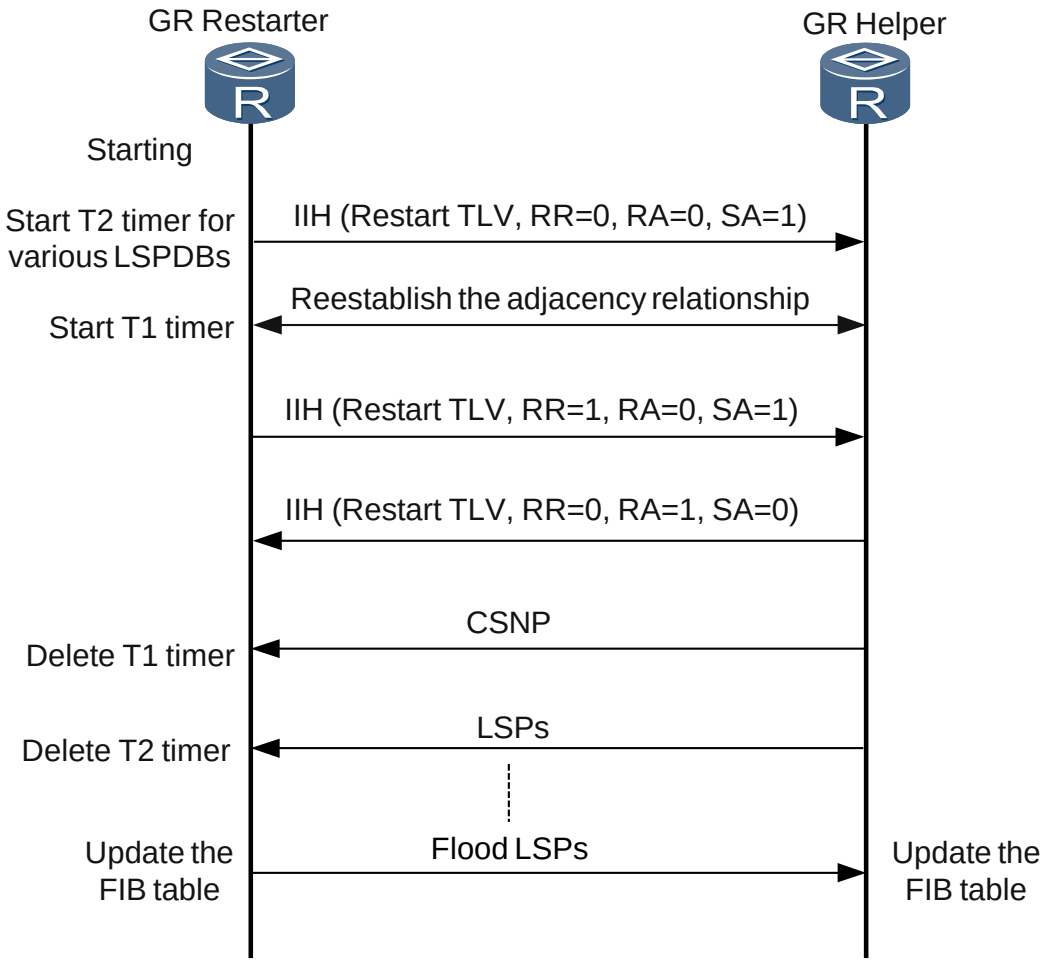
If the GR helper does not support GR?, it ignores the restart TLV and resets the adjacency with the GR restarter according to the normal processing of IS-IS.

3. After the GR restarter receives the IIH response packet, in which RR is set to 0 and RA is set to 1, from the neighbor, it performs the following actions:
   - Compares the current value of the T3 timer with the value of the Remaining Time field in the packet. The smaller value is taken as the value of the T3 timer.
   - Deletes the T1 timer maintained by the interface that receives the ACK packet and CSNPs.
   - If the interface does not receive the ACK packet or CSNPs, the GR restarter constantly resets the T1 timer and resends the IIH packet that contains the restart TLV. If the number of timeouts of the T1 timer exceeds the threshold value, the GR restarter forcibly deletes the T1 timer and turns to the normal IS-IS processing to complete LSDB synchronization.
4. After the GR restarter deletes the T1 timers on all interfaces, the synchronization with all neighbors is complete when the CSNP list is cleared and all LSPs are collected. The T2 timer is then deleted.
5. After the T2 timer is deleted, the LSDB of the level is synchronized.
   - In the case of a Level-1 or Level-2 router, SPF calculation is triggered.
   - In the case of a Level-1-2 router, ?determine whether the T2 timer on the router of the other level is also deleted. If both T2 timers are deleted, SPF calculation is triggered. Otherwise, the router waits for the T2 timer of the other level to expire.
6. After all T2 timers are deleted, the GR restarter deletes the T3 timer and updates the FIB table. The GR restarter re-generates the LSPs of each level and floods them. During LSDB synchronization, the GR restarter deletes the LSPs generated before restarting.
7. At this point, the IS-IS restarting of the GR restarter is complete.

IS-IS Starting

The starting device does not retain the FIB table. The starting device depends on the neighbors, whose adjacency with itself is Up before it starts, to reset their adjacency and suppress the neighbors from advertising their adjacency. The IS-IS starting process is different from the IS-IS restarting process, as shown in Figure 16-43.

Figure 16-43 IS-IS starting



1. After the GR restarter is started, it performs the following actions:
   - Starts the T2 timer for the synchronization of LSDBs of each level.
   - Sends IIH packets that contain the restart TLV from all interfaces.

     If RR in the packet is set to 0, a router is started.

     If SA in the packet is set to 1, the router requests its neighbor to suppress the advertisement of their adjacency before the neighbor receives the IIH packet in which SA is set to 0.
2. After the neighbor receives the IIH packet that carries the restart TLV, it performs the following actions depending on whether GR is supported:
   - GR is supported.

     Re-initiates the adjacency.

     Deletes the description of the adjacency with the GR restarter from the sent LSP. The neighbor also ignores the link connected to the GR restarter when performing SPF calculation until it receives an IIH packet in which SA is set to 0.
   - GR is not supported.

     Ignores the restart TLV and resets the adjacency with the GR restarter.

Replies with an IIH packet that does not contain the restart TLV. The neighbor then returns to normal IS-IS processing. In this case, the neighbor does not suppress the advertisement of the adjacency with the GR restarter. On a P2P link, the neighbor also sends a CSNP.

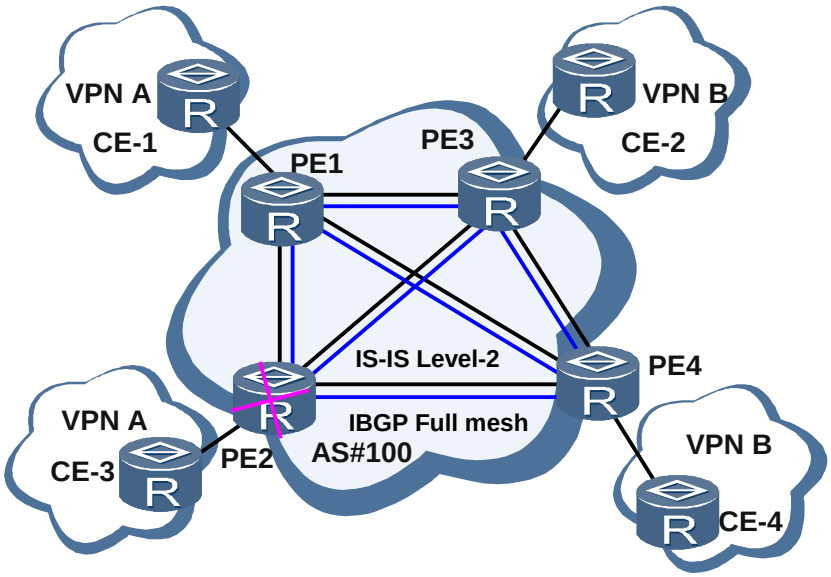
3. After the adjacency is re-initiated, the GR restarter re-establishes the adjacency with the neighbors on each interface. When an adjacency set on an interface is in the Up state, the GR restarter starts the T1 timer for the interface.
4. After the T1 timer expires, the GR restarter sends an IIH packet in which both RR and SA are set to 1.
5. After the neighbor receives the IIH packet, it replies with an IIH packet, in which RR is set to 0 and RA is set to 1, and sends a CSNP.
6. After the GR restarter receives the IIH ACK packet and CSNP from the neighbor, it deletes the T1 timer.

   If the GR restarter does not receive the IIH packet or CSNP, it constantly resets the T1 timer and resends the IIH packet in which RR and SA are set to 1. If the number of the timeouts of the T1 timer exceeds the threshold value, the GR restarter forcibly deletes the T1 timer and turns to the normal IS-IS processing to complete LSDB synchronization.
7. After receiving the CSNP from the helper, the GR restarter synchronizes the LSDB.
8. After the LSDB of this level is synchronized, the T2 timer is deleted.
9. After all T2 timers are deleted, the SPF calculation is started and LSPs are regenerated and flooded.
10. At this point, the IS-IS starting of the GR restarter is complete.

Application Environment

GR is typically applied on PEs, especially single point PEs. To prevent the scenario where a single point PE fails or master/slave switchover occurs on a PE due to maintenance operations, such as upgrading the software version, GR is configured to ensure non-stop forwarding of key services. Figure 16-44 shows the networking for the application of GR.

Figure 16-44 GR on provider network



NOTE

NSF is deployed on PE2 to prevent single-node failure on PE2. (IS-IS GR, and LDP GR run on PE2).

P2P network is recommended when IS-IS GR is enabled. The **isis circuit-type p2p** command sets the network type of an interface to emulate a P2P interface.

On the PEs, IS-IS, or LDP GR is run. On the Ps, IS-IS or LDP GR is run. The MPU/SRUs on the PEs and Ps work in backup mode.

16.8.13 IS-IS Wide Metric

A small range of metrics cannot meet the requirements of large-scale networks.

In the earlier ISO 10589, the greatest value of an interface metric was only 63. TLV type 128 and TLV type 130 contained information about routes; TLV type 2 contained information about IS-IS neighbors.

As defined in RFC 3784, the value of an interface metric can be extended to 16777215, and the metric of a route can reach 4261412864. With IS-IS wide metric enabled, TLV type 135 contains information about routes; TLV type 22 contains information about IS-IS neighbors.

- The following TLVs are used in narrow mode:
  - IP Internal Reachability: carries internal routes.
  - IP External Reachability: carries external routes.
  - IS Neighbors: carries information about neighbors.
- The following TLVs are used in wide mode:
  - Extended IP Reachability: replaces the earlier IP reachability TLV and carries information about routes. This TLV expands the range of route cost to 4 bytes and carries sub-TLVs.
  - IS Extended Neighbors: carries information about neighbors.

NOTE
IS-IS in wide mode and IS-IS in narrow mode cannot communicate. If IS-IS in wide mode and IS-IS in narrow mode need to communicate, you must change the mode to enable all routers on the network to receive packets sent by other routers.

Table 16-14 Receiving and sending modes

Mode\Receiving and Sending	Receiving	Sending
narrow	narrow	narrow
narrow-compatible	narrow&wide	narrow
compatible	narrow&wide	narrow&wide
wide-compatible	narrow&wide	wide
wide	wide	wide

When the cost-style is set to compatible, IS-IS sends the information in narrow mode and then in wide mode.

Process

NOTICE
A cost-style change causes the IS-IS process to restart. Be cautious in your use of the **cost-style** command.

- Changing the sending mode from narrow to wide

  The information that used to be carried by TLV type 128, TLV type 130, and TLV type 2 is now carried by TLV type 135 and TLV type 22.
- Changing the sending mode from wide to narrow

  The information that used to be carried by TLV type 135 and TLV type 22 is now carried by TLV type 128, TLV type 130, and TLV type 2.
- Changing the sending mode from narrow/wide to narrow&wide

  The information that used to be carried in narrow/wide mode is now carried by TLV type 128, TLV type 130, TLV type 2, TLV type 135, and TLV type 22.

16.8.14 BFD for IS-IS

BFD functions as a simple "Hello" protocol. In many aspects, it is similar to the adjacency test of a routing protocol.

Two systems periodically send BFD packets on the path between them. If one system does not receive any BFD packets from its peer within the detection period, the system detects that the bidirectional path to its peer is faulty. Under some conditions, systems need to negotiate the sending and receiving rates to reduce the load.

BFD is classified into static BFD and dynamic BFD.

NOTE
BFD uses the local discriminator and remote discriminator to differentiate multiple BFD sessions between the same pair of systems.

- Static BFD

  In static BFD, BFD session parameters including local and remote discriminators are set using commands, and the requests for establishing BFD sessions are manually delivered.

- Dynamic BFD(including BFD for IPv4)

  In dynamic BFD, the establishment of BFD sessions is triggered by routing protocols. The local discriminator is dynamically assigned, and the remote discriminator is learned by a routing protocol.

In BFD for IS-IS, the establishment of a BFD session is dynamically triggered by IS-IS instead of being performed manually by an administrator. When detecting a fault, BFD notifies IS-IS of the fault through the RM module. IS-IS then sets the status of the associated neighbor relationship to Down, immediately advertises the changed Link State PDU (LSP), and performs incremental SPF. In this manner, fast route convergence is implemented.

Generally, the interval for sending Hello packets is set to 10s. The interval for advertising that a neighbor is Down, that is, the Holddown time for keeping the neighbor relationship, is three times the interval for sending Hello packets. If a router does not receive any Hello packet from its neighbor within the Holddown time, the router deletes the associated neighbor relationship.

A router can detect a neighbor fault at only the second level. As a result, a large number of packets may be lost on a high-speed network.

To solve the problem, BFD provides link fault detection featuring light load and high speed (in milliseconds).

BFD can provide millisecond-level fault detection. BFD does not take the place of the Hello mechanism of IS-IS, but works with IS-IS to more quickly detect the faults that occur on neighboring devices or links, and instructs IS-IS to recalculate routes to correctly guide packet forwarding.

Static BFD

In static BFD, BFD session parameters including local and remote discriminators are set using commands, and the requests for establishing BFD sessions are manually delivered.

In this mode, the creation and deletion of BFD sessions also need to be triggered manually, which is inflexible and configuration errors can occur from user mistakes. For example, the local discriminator and remote discriminator are incorrectly configured, which causes abnormal functioning of the BFD session.

Dynamic BFD

Dynamic BFD is more flexible than static BFD. In dynamic BFD, routing protocols trigger the establishment of BFD session. The establishment of a BFD-for-IPv4 session is triggered by IS-IS when an IPv4 neighbor relationship is set up.

In setting up a new neighbor relationship, IS-IS sends parameters of the neighbors and detection parameters (including source and destination IP addresses) to BFD. BFD then sets up a session according to the received parameters. Dynamic BFD is more flexible than static BFD.

The RM module provides related services for association with the BFD module for IS-IS. Through RM, IS-IS prompts BFD to set up or tear down BFD sessions by sending notification messages. In addition, BFD events are transmitted to IS-IS through RM.

Establishment and Deletion of BFD Sessions

- Conditions for setting up a BFD session
  - Basic IS-IS functions are configured on each router and IS-IS is enabled on the interfaces of the routers.
  - BFD is enabled on each router, and BFD for IPv4 is enabled on interfaces or processes of the routers.
  - BFD for IPv4 is enabled on interfaces or processes, and the status of the neighboring router is Up (the DIS must be elected on a broadcast network).
- Process of setting up a BFD session
  - P2P network

    After the conditions for setting up a BFD session are satisfied, IS-IS instructs BFD through RM to directly set up a BFD session between neighbors.
  - Broadcast network

    After the conditions for establishing BFD sessions are met, and the DIS is elected, IS-IS instructs BFD through RM to establish a BFD session between the DIS and each router. No BFD session is established between non-DISs.

  On a broadcast network, the routers (including non-DIS routers) of the same level on the same network segment can set up neighbor relationships. In the implementation of IS-IS BFD, however, BFD sessions are set up between the DIS and non-DIS devices rather than between non-DISs. On a P2P network, BFD sessions are directly set up between neighbors.

  If a Level-1-2 neighbor relationship is set up between two routers on a link, IS-IS sets up two BFD sessions for the Level-1 and Level-2 neighbors on a broadcast network, but sets up only one BFD session on a P2P network.
- Conditions for tearing down a BFD session
  - P2P network

    When a neighbor relationship that was set up on P2P interfaces by IS-IS is down (that is, the neighbor relationship is not in the Up state) or when the IP protocol type of a neighbor is deleted, IS-IS tears down the BFD session.
  - Broadcast network

    When a neighbor relationship that was set up on P2P interfaces by IS-IS is torn down (that is, the neighbor relationship is not in the Up state) when the IP protocol type of a neighbor is deleted, or when the DIS is re-elected, IS-IS tears down the BFD session.

  When the configurations of a dynamically established BFD session are deleted or BFD for IS-IS is disabled on an interface, all BFD sessions to which neighbor relationships on the interface correspond-between devices or between devices and the DIS are deleted.

  After dynamic BFD is globally disabled in an IS-IS process, the BFD sessions on all the interfaces in this IS-IS process are deleted.

  NOTE

  BFD detects only one-hop links between IS-IS neighbors, because IS-IS establishes only one-hop neighbor relationships.
- Response to the Down event of a BFD session

When detecting a link failure, BFD generates a Down event, and then notifies RM of the event. RM then instructs IS-IS to deletes the neighbor relationship. IS-IS recalculates routes to speed up route convergence on the entire network. After BFD for IPv4 informs IS-IS of the link failure, IS-IS changes only the IPv4 route.

When a router and its neighbor are Level-1-2 routers, they set up two neighbor relationships, that is, the Level-1 neighbor relationship and the Level-2 neighbor relationship. Then, IS-IS sets up two BFD sessions for the Level-1 neighbor relationship and Level-2 neighbor relationship. In this case, the RM module deletes the neighbor relationship of a specific level.

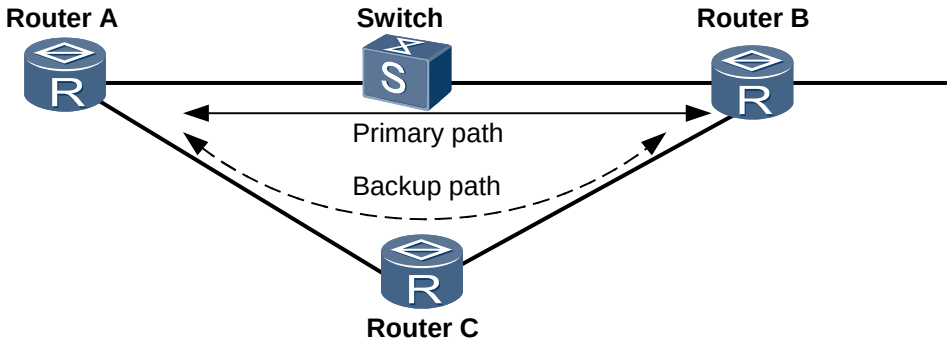
Applicable Environment

NOTICE

You must configure BFD according to the actual network environment. If timer parameters are set improperly, network flapping may occur.

BFD for IS-IS can quickly sense link changes to implement fast route convergence.

Figure 16-45 Networking for BFD for IS-IS



BFD for IS-IS configuration requirements are as follows:

- Enable IS-IS on the routers, as shown in Figure 16-45.
- Enable BFD globally.
- Enable BFD for IS-IS on Router A and Router B.

When the link between Router A and Router B becomes faulty, BFD can quickly detect the fault and notify IS-IS. IS-IS then changes the neighbor relationship on the interface to Down and deletes the IP protocol type to which the neighbor relationship corresponds, which triggers route calculation. In addition, IS-IS updates LSPs so that neighbors such as Router C can receive updated LSPs from Router B.

16.8.15 IS-IS Authentication

To ensure network security, IS-IS authentication encrypts IS-IS packets by adding the authentication field to packets to ensure network security. When a local router receives IS-IS

packets from a remote router, the local router discards the packets if the authentication passwords do not match. This protects the local router.

Based on the types of packets, the authentication is classified as follows:

- Area authentication: is configured in the IS-IS process view to authenticate Level-1 CSNPs, PSNPs, and LSPs.
- Routing domain authentication: is configured in the IS-IS process view to authenticate Level-2 CSNPS, PSNPs, and LSPs.
- Interface authentication: is configured in the interface view to authenticate Level-1 and Level-2 Hello packets.

Based on the authentication modes of packets, authentication is classified into the following types:

- Plain text authentication: is a simple authentication mode in which passwords are directly added to packets.
- MD5 authentication: uses the MD5 algorithm to encrypt passwords before they are added to packets, which improves password security.

NOTE
MD5 authentication is commended.

IS-IS provides a TLV to carry authentication information. The TLV components are as follows:

- Type (of the authentication packets): is defined by ISO as 10, with a length of 1 byte.
- Length: indicates the length of the authentication TLV, which is 1 byte.
- Value: indicates the contents of the authentication, including authentication type and authenticated password, which ranges from 1 to 254 bytes.
  - Type 0 is reserved.
  - Type 1 indicates plain text authentication.
  - Type 54 indicates MD5 authentication.
  - Type 255 is used to route domain private authentication methods.

The authentication password is saved in the following modes:

- The authentication password for IIH packets is saved on interfaces. It is implemented as interface authentication.
- The authentication password for Level-1 LSPs and SNPs are saved in the IS-IS process. It is implemented as area authentication.
- The authentication password for Level-2 LSPs and SNPs are saved in the IS-IS process. It is implemented as routing domain authentication.

Interface authentication can be classified as follows:

- A router sends authentication packets with the authentication TLV and verifies the authentication information of the packets it receives.
- A router sends authentication packets with the authentication TLV but does not verify the authentication information of the packets it receives.

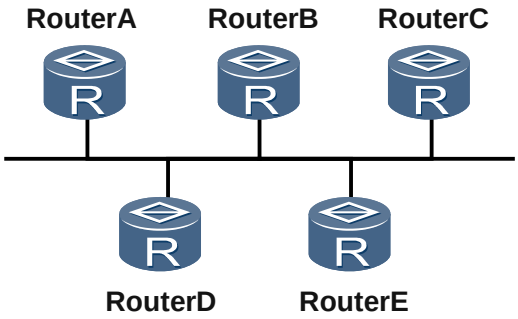
For area authentication and routing domain authentication, you can set a router to authenticate SNPs and LSPs separately in the following ways:

- A router sends LSPs and SNPs that carry the authentication TLV and verifies the authentication information of the LSPs and SNPs it receives.

- A router sends LSPs that carry the authentication TLV and verifies the authentication information of the LSPs it receives. The router sends SNPs that carry the authentication TLV but does not verify the authentication information of the SNPs it receives.
- A router sends LSPs that carry the authentication TLV and verifies the authentication information of the LSPs it receives. The router sends SNPs without the authentication TLV and does not verify the authentication information of the SNPs it receives.
- A router sends LSPs and SNPs that carry the authentication TLV but does not verify the authentication information of the LSPs and SNPs it receives.

Application Environment

Figure 16-46 Networking for IS-IS authentication on a broadcast network



The requirements for IS-IS authentication on a broadcast network are as follows:

- IS-IS neighbor relationships can be set up between multiple routers on the same network only when interface authentication is configured in the same manner on all the routers.
- When multiple routers are in the same area, you must configure area authentication the same way on all the routers to ensure synchronization of their Level-1 LSDBs.
- When Level-2 neighbor relationships are set up between multiple routers, you must configure routing domain authentication the same way on all the routers to ensure the synchronization of their Level-2 LSDBs.

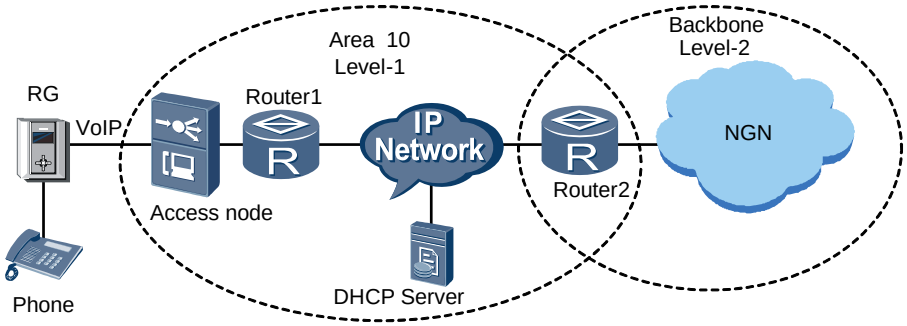
16.8.16 Configuration Example of IS-IS

This operation enables the corresponding device configured data to run the IS-IS protocol on the access device.

Service Requirements

- The access device forwards the access VoIP service through the Layer 3 interface to the NGN network.
- The access device obtains the routes of the NGN networking through the IS-IS protocol. The area ID of the Level-2 router differs from the area ID of the Level-1-2 router to which the Level-2 router connects.

Figure 16-47 Example network for configuring IS-IS



Data Plan

Table 16-15 provides the data plan for configuring IS-IS.

Table 16-15 Data plan for configuring IS-IS

Item	Data
Access node	IS-IS process ID: 1
	NET (Network entity title): 10.0000.0000.0001.00, where: • Area ID: 10 • System ID: 0000.0000.0001 • Level: Level-1 • Host name: MA5600T
	IS-IS interface: • Port number: 0/20/0 • VLAN ID: 20 • IP address: 5.5.5.5/16
Router1	IS-IS process ID: 1
	NET (Network entity title): 10.0000.0000.0002.00, where: • Area ID: 10 • System ID: 0000.0000.0002 • Level: Level-1 • Host name: Router1
	IS-IS interface: 1/0/0 IP address: 8.8.8.8/16
Router2	IS-IS process ID: 1
	NET (Network entity title): 10.0000.0000.0005.00, where: • Area ID: 10 • System ID: 0000.0000.0005

Item	Data
	• Level: Level-1-2 • Host name: Router2
	IS-IS interface: 1/0/0 IP address: 9.9.9.9/16

Procedure

- Configure IS-IS on the access node.
  a. Configure the Layer 3 interface.

```
huawei(config)#vlan 20 standard
huawei(config)#port vlan 20 0/20 0
huawei(config)#interface vlanif 20
huawei(config-if-vlanif20)#ip address 5.5.5.5 16
huawei(config-if-vlanif20)#quit
```

  b. Start the IS-IS process.

```
huawei(config)#isis 1
huawei(config-isis-1)#
```

  c. Configure the NET.

```
huawei(config-isis-1)#network-entity 10.0000.0000.0001.00
```

  d. Configure the router level.

```
huawei(config-isis-1)#is-level level-1
```

  e. Configure the local host name.

```
huawei(config-isis-1)#is-name MA5600T
huawei(config-isis-1)#quit
```

  f. Enable the IS-IS function on an interface.

```
huawei(config)#interface vlanif 20
huawei(config-if-vlanif20)#isis enable 1
```

- Configure IS-IS on Router1.

  The process of configuring IS-IS on Router1 is similar to that of configuring IS-IS on the access node. The details are not provided in this chapter.

- Configure IS-IS on Router2.

  The process of configuring IS-IS on Router2 is similar to that of configuring IS-IS on the access node. The details are not provided in this chapter.

----End

Result

- Run the **display isis lsdb** command and you can query the IS-IS LSDB.

- Run the **display isis route** command and you can query the IS-IS route. The routing table of the Level-1 router should have a default route, and the next hop should be the Level-1-2 router. The Level-2 router should have the routes to all the Level-1 routers and the Level-2 routers.

Configuration File

```
vlan 20 standard
port vlan 20 0/20 0
interface vlanif 20
ip address 5.5.5.5 16
quit
isis 1
network-entity 10.0000.0000.0001.00
is-level level-1
is-name MA5600T
quit
interface vlanif 20
isis enable 1
```

16.8.17 References

Table 16-16 The following table lists the references.

Document No.	Document Name	Protocol Compliance
ISO 10589	IS-IS intra-domain routing protocol	Fully compliant
RFC 1142	OSI IS-IS Intra-domain Routing Protocol	Fully compliant
RFC 1195	Use of OSI IS-IS for Routing in TCP/IP and Dual Environments	Fully compliant
RFC 2104	HMAC: Keyed-Hashing for Message Authentication	Fully compliant
RFC 2763	Dynamic Hostname Exchange Mechanism for IS-IS	Fully compliant
RFC 2966	Domain-wide Prefix Distribution with Two-Level IS-IS	Fully compliant
RFC 2973	IS-IS Mesh Groups	Fully compliant
RFC 3277	IS-IS Transient Blackhole Avoidance	Fully compliant
RFC 3567	Intermediate System to Intermediate System (IS-IS) Cryptographic Authentication	Fully compliant
RFC 3719	Recommendations for Interoperable Networks using IS-IS	Fully compliant
RFC 3784	IS-IS Extensions for Traffic Engineering	Fully compliant

Document No.	Document Name	Protocol Compliance
RFC 3786	Extending the Number of IS-IS LSP Fragments Beyond the 256 Limit	Fully compliant
RFC 3787	Recommendations for Interoperable IP Networks using IS-IS	Fully compliant
RFC 4444	Management Information Base for IS-IS	Fully compliant
RFC 5130	A Policy Control Mechanism in IS-IS Using Administrate	Fully compliant
RFC 5301	Dynamic Hostname Exchange Mechanism for IS-IS	Fully compliant
RFC 5302	Domain-wide Prefix Distribution with Two-Level IS-IS	Fully compliant
RFC 5305	IS-IS Extensions for Traffic Engineering	Fully compliant
RFC 5306	Restart Signaling for IS-IS	Fully compliant
RFC 5308	Routing IPv6 with IS-IS	Fully compliant
RFC 5309	Point-to-Point Operation over LAN in Link State Routing Protocols	Fully compliant
RFC 5311	Simplified Extension of Link State PDU (LSP) Space for IS-IS	Fully compliant
RFC 6232	Purge Originator Identification TLV for IS-IS	Partially compliant Purge Originator Identification (POI) TLVs defined in RFC 6232 can be received but not generated.
draft-ietf-isis-mi-06	IS-IS Multi-Instance	Partially compliant IS-IS multi-instance is supported, but an interface can be bound to only one instance.
draft-ietf-isis-wg-multi-topology-11	M-ISIS: Multi Topology (MT) Routing in IS-IS	Fully compliant

16.9 OSPF

Open Shortest Path First (OSPF) is an interior gateway protocol (IGP) based on the link state developed by the Internet Engineering Task Force (IETF).

16.9.1 Introduction to OSPF

Definition

The Open Shortest Path First (OSPF) protocol, developed by the Internet Engineering Task Force (IETF), is a link-state Interior Gateway Protocol (IGP).

Currently, OSPF Version 2 as defined in RFC 2328, is intended for IPv4. Defined in RFC 2740, OSPF Version 3 is intended for IPv6. In this document, the term OSPF refers to OSPFv2, unless otherwise stated.

Purpose

Before the emergence of OSPF, the Routing Information Protocol (RIP) was widely used as an IGP on networks.

RIP is a routing protocol based on the distance vector algorithm. Due to its slow convergence, routing loops, and poor scalability, RIP has been gradually replaced by OSPF.

As a link-state protocol, OSPF can solve many problems inherent in RIP. OSPF has the following advantages:

- Transmits packets in multicast mode to reduce load on routers that do not run OSPF.
- Supports Classless Inter-domain Routing (CIDR).
- Supports load balancing among equal-cost routes.
- Supports packet encryption.

OSPF is now widely accepted for use as an IGP.

16.9.2 Fundamentals of OSPF

OSPF has the following advantages:

- Divides an Autonomous System (AS) into a single area or multiple logical areas.
- Sends Link State Advertisements (LSA) to advertise routes.
- Synchronizes routing information by exchanging OSPF packets between routers in OSPF areas.
- Encapsulates the OSPF packets in IP packets and sends the packets in unicast or multicast mode.
- Enabling the feature on an OSPF interface is supported to allow users to manage OSPF using NMS.
- The same-router-ID detection and recovery function is supported. After OSPF detects same router IDs, it selects a new router ID to avoid route flapping.

OSPF Packet Type

Table 16-17 OSPF packet types

Packet Type	Function
Hello	Hello packets are sent periodically to discover and maintain OSPF neighbor relationships.
Database Description (DD)	DD packets carry brief information about the local Link State Database (LSDB) and are used to synchronize the LSDBs of two routers.
Link State Request (LSR)	LSR packets are used to request the required LSAs from neighbors. LSR packets are sent only after DD packets have been exchanged successfully.
Link State Update (LSU)	LSU packets are used to send the required LSAs to neighbors.
Link State Acknowledgment (LSAck)	LSAck packets are used to acknowledge the received LSAs.

LSA Type

Table 16-18 OSPF LSA types

LSA	Function
Router-LSA (Type1)	Describes the link status and link cost of the MA5600T/MA5603T/MA5608T. Generated by each MA5600T/MA5603T/MA5608T and advertised in the area to which the MA5600T/MA5603T/MA5608T belongs.
Network-LSA (Type2)	Describes the link status of all routers in the local network segment. Generated by a designated router (DR) and advertised in the area to which the DR belongs.
Network-summary-LSA (Type3)	Describes the routes in a network segment and advertises the routes to the related non-totally STUB or NSSA area.
ASBR-summary-LSA (Type4)	Describes routes to an Autonomous System Boundary Router (ASBR). Generated by an ABR and advertised in the related areas, except the area to which the ASBR belongs.
AS-external-LSA (Type5)	Describes routes to a destination outside the AS. Generated by an ASBR and advertised in all areas, except stub areas and Not-So-Stubby Areas (NSSA).
NSSA-LSA (Type7)	Describes routes to a destination outside the AS. Generated by an ASBR and advertised in NSSAs only.
Opaque-LSA (Type9/Type10/Type11)	Provides a general mechanism for OSPF extension: • Type9 LSAs are advertised in the network segment

LSA	Function
	where interfaces reside. Graceful LSAs used to support GR are one example of Type9 LSAs. • Type10 LSAs are advertised in an area. LSAs used to support TE are one example of Type10 LSAs. • Type11 LSAs are advertised in an AS. Currently, no application examples of Type11 LSAs exist.

Router Type

Figure 16-48 illustrates the common types of routers in OSPF.

Figure 16-48 Router types

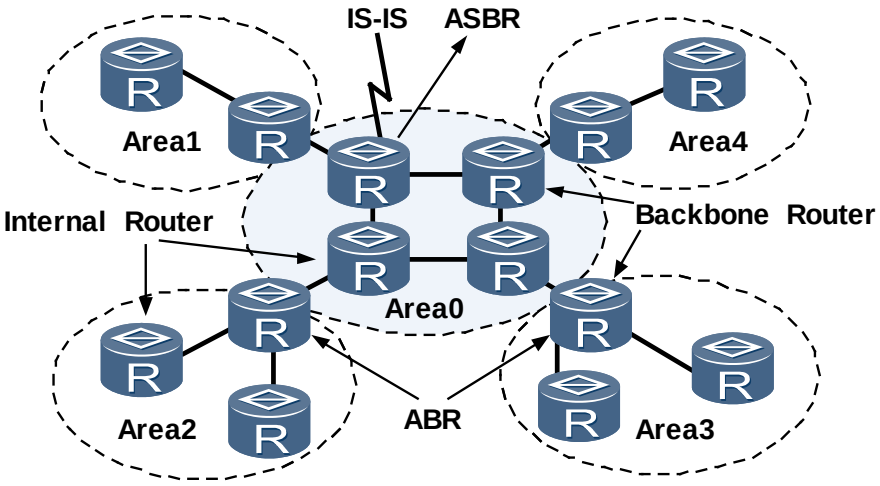


Table 16-19 OSPF router types

Router	Description
Internal router	All interfaces of an internal router belong to the same OSPF area.
Area Border Router (ABR)	An ABR can belong to two or more areas; one of the areas must be a backbone area. An ABR is used to connect the backbone area and non-backbone areas. It can be physically or logically connected to the backbone area.
Backbone router	At least one interface on a backbone router belongs to the backbone area. All ABRs and internal routers in Area 0 are backbone routers.
AS Boundary Router (ASBR)	An ASBR exchanges routing information with other ASs. An ASBR does not have to reside at the boundary of an

Router	Description
	AS. It can be an internal router or an ABR.

OSPF Route Type

Inter-area routes and intra-area routes define the network structure of an AS. External routes define how to select a route to a destination outside an AS. OSPF classifies the imported AS external routes into Type1 and Type2 external routes.

Table 16-20 lists route types in descending order of priority.

Table 16-20 OSPF route types

Route	Description
Intra area	Intra-area routes
Inter area	Inter-area routes
Type1 external route	Because of the high reliability of Type1 external routes, the calculated cost of external routes equals that of AS internal routes, and. In other words, the cost of a Type1 external route equals the cost of the route from the router to the corresponding ASBR plus the cost of the route from the ASBR to the destination.
Type2 external route	Because of the low reliability of Type2 external routes, their costs are considered to be greater than the cost of any internal path to an ASBR. The cost of a Type2 external route equals the cost of the route from the ASBR to the destination.

Area Type

Table 16-21 OSPF area types

Area	Function
Totally stub area	Allows Type3 default routes that are advertised by an ABR, and denies inter-area routes and the routes outside an AS.
Stub area	Allows inter-area routes, unlike a totally stub area.
NSSA area	Imports routes from outside an AS, unlike a stub area. An ASBR advertises Type7 LSAs in the local area.
Totally NSSA	Denies inter-area routes, unlike an NSSA.

OSPF Network Type

OSPF classifies network, in terms of link layer protocols, into the following types listed in Table 16-22.

Table 16-22 OSPF network types

Network	Description
Broadcast	If the link layer protocol is Ethernet or Fiber Distributed Data Interface (FDDI), OSPF defaults the network type to broadcast. In broadcast networks: • Hello and LSAck packets are transmitted in multicast mode. LSU packets are first transmitted in multicast mode and retransmitted in unicast mode. The address 224.0.0.5 is the reserved IP multicast address of the OSPF router, and the address 224.0.0.6 is the reserved IP multicast address of the OSPF DR. • DD packets and LSR packets are transmitted in unicast mode.
Non-Broadcast Multiple Access (NBMA)	If the link layer protocol is ATM, OSPF defaults the network type to NBMA. In NBMA networks, protocol packets, such as Hello, DD, LSR, LSU, and LSAck packets, are transmitted in unicast mode.
Point-to-Multipoint (P2MP)	Regardless of the link layer protocol, OSPF does not default the network type to P2MP. A P2MP network must be forcibly changed from other network types. The common practice is to change a non-fully connected NBMA network to a P2MP network. In P2MP networks: • Hello packets are transmitted in multicast mode through the multicast address 224.0.0.5. • Other protocol packets, such as DD, LSR, LSU, and LSAck packets, are transmitted in unicast mode.
Point-to-point (P2P)	If the link layer protocol is PPP, HDLC, or LAPB, OSPF defaults the network type to P2P. In broadcast networks: • In P2P networks, protocol packets, such as Hello, DD, LSR, LSU, and LSAck packets, are transmitted in multicast mode through the multicast address 224.0.0.5. • LSU packets are retransmitted in unicast mode.

Stub Area

A stub area is a special area where ABRs do not flood the received external routes. In a stub area, the size of the routing table of routers and routing information in transmission are greatly reduced.

Configuring a stub area in a network is optional. Not all areas can be configured as stub areas. Generally, a stub area is a non-backbone area with only one ABR and is located at the AS boundary.

To ensure the reachability of a destination outside an AS, the ABR in a stub area generates a default route and advertises it to non-ABRs in the stub area.

When you configure a stub area, note the following:

- The backbone area cannot be configured as a stub area.
- If an area needs to be configured as a stub area, use the **stub** command to configure all the routers in this area.
- An ASBR cannot exist in a stub area. That is, external routes are not flooded in the stub area.
- A virtual link cannot pass through a stub area.

OSPF Route Aggregation

Route aggregation occurs when routes with the same prefix are aggregated into one route and the aggregated route is advertised in other areas.

After route aggregation, the route information can be reduced. Consequently, the size of routing tables is reduced, which improves the performance of routers.

Route aggregation can be carried out in the following ways:

- ABR aggregation

  When an ABR transmits routing information to other areas, the router originates Type3 LSAs per network segment. If any consecutive segments exist in this area, you can run the related command to aggregate these segments into one segment. An ABR sends only one aggregated LSA. Any LSA that belongs to the aggregated network segment specified by the command is not transmitted separately.
- ASBR aggregation

  After route aggregation is enabled, if the local router is an ASBR, it aggregates the imported Type5 LSAs within the aggregated address range. After an NSSA is configured, the ASBR aggregates the imported Type7 LSAs within the aggregated address range.

  If the local router is both an ABR and ASBR, it aggregates the Type5 LSAs that are transformed from Type7 LSAs.

OSPF Default Route

A *default route* is the route whose destination address and mask are all 0s. When a router does not have exact matching routes, it can forward packets through default routes.

OSPF default routes are applicable to the following situations:

- An ABR advertises the default Type3 summary-LSAs to instruct intra-area routers to forward packets to other areas.
- An ASBR advertises default Type5 ASE LSAs or Type7 NSSA LSAs to instruct intra-AS routers to forward packets to other ASs.

The principles for advertising OSPF LSAs that describe default routes are as follows:

- An OSPF router advertises an LSA that describes a default route only when an interface on the OSPF router is connected to a network outside an area.

- If an OSPF router has already advertised an LSA that describes a default route, the OSPF route no longer learns LSAs of the same type advertised by other routers. The OSPF router calculates routes by using an LSA that describes a default route in an LSDB, but not an LSA of the same type advertised by another router.
- If the OSPF router needs to advertise an LSA that describes a default route only with the help of another route, the route cannot be the one in the local routing domain. That is, it cannot be the one learned by the local OSPF process. The external default route guides forwarding outside the local OSPF routing domain, but the next hop of the routes in the local OSPF routing domain are inside the local OSPF routing domain, and fails to forward packets outside the local OSPF routing domain.
- The router checks whether there is any peer with the state of **full** in area 0 before advertising the default route. The router advertises the default route only when there are such peers because if there is no such peer, the backbone area cannot forward packets and advertising the default route is meaningless.

Table 16-23 shows the advertisement of default routes in different areas.

Table 16-23 Principles for advertising area-specific default routes

Area Type	Advertising Principles
Common area	By default, no default route is generated in a common area, even if a default route exists in the common area. After a default route is generated by another process, the default route must be advertised within an entire OSPF AS. To help OSPF generate a default route, you need to run a command on an ASBR. After the configuration, a default ASE LSA (Type5 LSA) is generated and advertised in the entire OSPF AS.
Stub area	AS external routes in Type5 LSAs cannot be advertised in a stub area. Routers in the stub area must learn AS external routes from an ABR. The ABR automatically generates a default summary-LSA (Type3 LSA) and advertises it in the entire stub area. Routers in the stub area can obtain reachable AS external routes through the ABR.
Totally stub area	AS external routes in Type5 LSAs or inter-area routes in Type3 LSAs cannot be advertised in a totally stub area. Routers in the totally stub area have to learn AS external routes and routes to other areas through an ABR. To help OSPF generate a default router, you need to configure a totally stub area. After the totally stub area is configured, an ABR automatically generates a default summary-LSA (Type3 LSA) and advertises it to the entire totally stub area. Routers in the totally stub area can obtain reachable AS external routes and routes to other areas through the ABR.
NSSA area	A small number of AS external routes that are obtained through the ASBR in the NSSA can be imported to an NSSA. Routes to other areas in ASE LSAs (Type5 LSAs) cannot be advertised in the NSSA. AS external routes are imported by the ASBR, and other external routes are advertised through other areas. The ABR generates a default NSSA LSA (Type7 LSA) automatically and advertises it in the entire NSSA. A small number of AS external routes can be obtained through the ASBR in the NSSA, and other

Area Type	Advertising Principles
	routes to other areas can be obtained through the ABR in the NSSA connected to ASBR in other areas. You need to run commands on the ASBR. The ASBR generates a default NSSA LSA (Type7 LSA) and advertises it to the entire NSSA. This way, external routes can be received through the ASBR in an NSSA. A Type7 LSA that describes a default route is neither translated into a Type5 LSA that describes a default route on an ABR nor advertised in the entire OSPF routing domain.
Totally NSSA area	External routes in ASE LSAs (Type5 LSAs) to other areas or inter-area routes in Type3 LSAs cannot be advertised in a totally NSSA. Routers in the totally NSSA learn routes to other areas from an ABR. You can configure a totally NSSA so that an ABR automatically generates a default Type7 LSA and advertises it to the entire totally NSSA. In this manner, routes to external areas and inter-area routes can be advertised in the totally NSSA through the ABR.

OSPF Route Filtering

By default, OSPF does not filter routes. OSPF supports the filtering of routes through routing policies.

OSPF routing policies include access control lists (ACLs), IP prefix lists, and route-policies. For details about these policies, see the section "Routing Policy"MA5600T/MA5603T/MA5608T *Feature Description - IP Routing*

OSPF route filtering is applicable in the following situations:

- Import of routes

  OSPF imports the routes that are learned by other protocols. You can configure routing policies to filter the routes so that OSPF imports only eligible routes.
- Advertisement of imported routes

  OSPF advertises the imported routes to neighbors.

  You can configure rules to filter the routing information to be advertised to neighbors. The filtering rules take effect only when being configured on ASBRs because only the ASBRs can import routes.
- Learning of routes

  You can configure rules, by which OSPF filters the received intra-area, inter-area, and AS external routes.

  The filtering action determines whether to add routing entries to the routing table. That is, only the routes that pass the filtering are added to the local routing table. All the routes, however, can still be advertised from the OSPF routing table.
- Learning of inter-area LSAs

  You can configure ABRs filter the incoming summary-LSAs of the local area using a command. This configuration takes effect only on ABRs, because only ABRs can advertise summary-LSAs.

- Advertisement of inter-area LSAs

  You can configure ABRs to filter the outgoing summary-LSAs of the local area through a command. This configuration takes effect only on ABRs.

Table 16-24 Differences between inter-area LSA learning and route learning

Inter-area LSA Learning	Route Learning
Filters the incoming LSAs of an area directly.	Filters only the calculated routes in LSAs to determine whether these routes are added to the local routing table.

OSPF Virtual Link

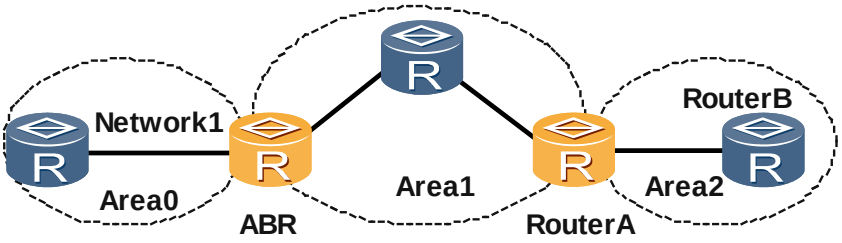
Virtual link refers to a logical channel established between two ABRs through a non-backbone area.

- A virtual link must be configured on both ends of the link; otherwise, it does not take effect.
- A transit area provides an internal route of a non-backbone area for both ends of the virtual link.

According to RFC 2328, during the deployment of OSPF, all non-backbone areas need to be connected to the backbone area. Otherwise, some areas will be unreachable.

As shown in Figure 16-49, Area 2 is not connected to the backbone area (Area 0), and Router A is not an ABR. Therefore, Router A does not advertise routing information of Network 1 in Area 0. As a result, Router B does not have the route to Network 1.

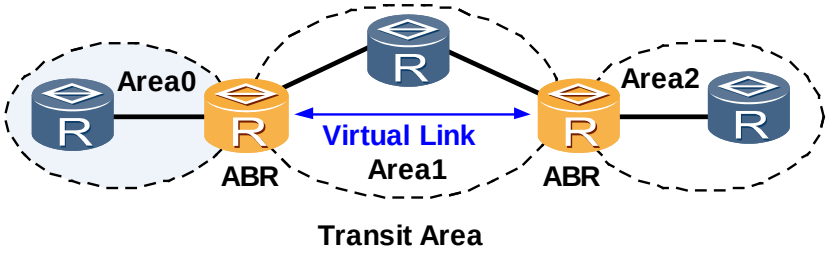
Figure 16-49 Non-Backbone area without connection to backbone area



Because of various limitations, in actual applications, physical connectivity between non-backbone areas and backbone areas cannot be ensured. To solve this problem, you can configure OSPF virtual links.

A virtual link is similar to a P2P connection between two ABRs. Similar to physical interfaces, you can configure the interfaces on both ends of the virtual link with parameters, such as the interval for sending Hello packets.

Figure 16-50 OSPF virtual link



As shown in Figure 16-50, OSPF packets transmitted between two ABRs are only forwarded by the OSPF routers that reside between the two ABRs. These routers detect that they are not the destinations of the packets, and forward the packets as common IP packets.

OSPF Multi-process

OSPF supports multi-processes. Multiple OSPF processes can run on the same router independently. Route interaction between different OSPF processes is similar to route interaction between different routing protocols.

An interface of a router belongs to only a certain OSPF process.

A typical application of OSPF multi-process is to run OSPF between PEs and CEs in the VPN where OSPF is also adopted in the backbone network. On the PEs, the two OSPF processes are independent of each other.

16.9.3 OSPF GR

Routers generally operate with separation of the control plane and forwarding plane. When the network topology remains stable, a restart of the control plane does not affect the forwarding plane, and the forwarding plane can still forward data properly. This separation ensures non-stop service forwarding.

In graceful restart (GR) mode, the forwarding plane continues to direct data forwarding after a restart occurs. The actions on the control plane, such as re-establishment of neighbor relationships and route calculation, do not affect the forwarding plane. Network reliability is improved because service interruption caused by route flapping is prevented.

Basic Concepts of OSPF GR

As mentioned in chapter 4, Graceful Restart (GR) is a technology used to ensure normal traffic forwarding and non-stop forwarding of key services during the restart of routing protocols.

Unless otherwise stated, GR described in this section refers to the GR technology defined in RFC 3623.

GR is one of the high availability (HA) technologies, which comprise a set of comprehensive technologies, such as fault-tolerant redundancy, link protection, faulty node recovery, and traffic engineering. As a fault-tolerant redundancy technology, GR is widely used to ensure non-stop forwarding of key services during master/slave switchover and system upgrade.

The following concepts are involved in GR:

- Grace-LSA

OSPF supports GR by flooding grace LSAs. Grace LSAs are used to inform the neighbor of the GR time, cause, and interface address when the GR starts and ends.

- Role of a router during GR
  - Restarter: is the router that restarts. The Restarter can be configured to support totally GR or partly GR.
  - Helper: is the router that helps the Restarter. The Helper can be configured to support planned GR or unplanned GR or to selectively support GR through the configured policies.
- Conditions that cause GR
  - Unknown: indicates that GR is triggered for an unknown reason.
  - Software restart: indicates that GR is triggered by commands.
  - Software reload/upgrade: indicates that GR is triggered by software restart or upgrade.
  - Switch to redundant control processor: indicates that GR is triggered by the abnormal master/slave switchover.
- GR period

  The GR period cannot exceed 1800 seconds. OSPF routers can exit from GR regardless of whether GR succeeds or fails, without waiting for GR to expire.

Classification of OSPF GR

Classification based on GR status:

- Totally GR: indicates that when a neighbor of a router does not support GR, the router exits from GR.
- Partly GR: indicates that when a neighbor does not support GR, only the interface associated with this neighbor exits from GR, whereas the other interfaces perform GR normally.

Classification based on the GR implementation mode:

- Planned GR: indicates that a router restarts or performs the master/slave switchover using a command. The Restarter sends a grace LSA before restart or master/slave switchover.
- Unplanned GR: indicates that a router restarts or performs the master/slave switchover because of faults. A router performs the master/slave switchover, without sending a grace LSA, and then enters GR after the slave board goes Up.

GR Process

- A router starts GR.

  In planned GR mode, after master/slave switchover is triggered through a command, the Restarter sends a grace LSA to all neighbors to notify them of the start, period, and cause of GR, and then performs the master/slave switchover.

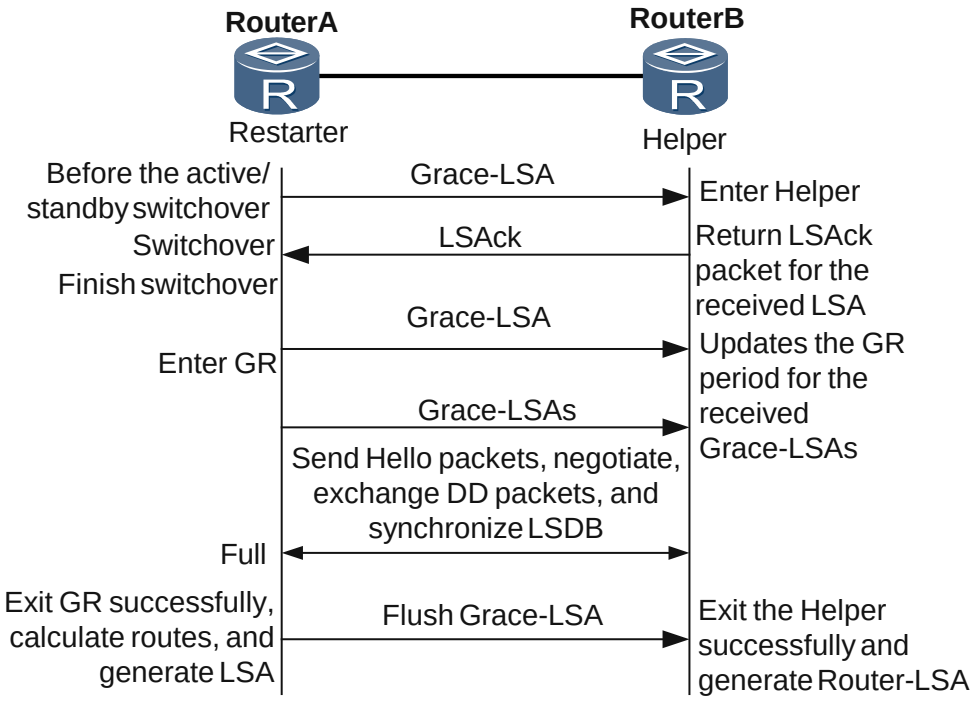
  In unplanned GR, the Restarter does not send the grace LSA.

  In unplanned GR mode, the Restarter sends a grace LSA immediately after the slave board goes Up, informing neighbors of the start, period, and cause of GR. The Restarter then sends a grace LSA to each neighbor five times consecutively. This ensures that neighbors receive the grace LSA. This operation is proposed by manufacturers but not defined by the OSPF protocol.

The Restarter sends a grace LSA to notify neighbors that it enters GR. During GR, neighbors keep neighbor relationships with the Restarter so that other routers cannot detect the switchover of the Restarter.

- The GR process runs, as shown in Figure 1

Figure 16-51 OSPF GR process



- The router exits from GR.

Table 16-25 Reasons that a router exits GR

Execution of GR	Restarter	Helper
GR succeeds.	Before GR expires, the Restarter re-establishes neighbor relationships with all neighbors before master/slave switchover.	After the Helper receives the grace LSA with the Age being 3600s from the Restarter, the neighbor relationship between the Helper and Restarter enters the Full state.
GR fails.	• GR expires, and neighbor relationships do not recover completely. • Router LSA or network LSA sent by the Helper causes Restarter to fail to perform bidirectional check. • Status of the interface that functions as the Restarter changes. • Restarter receives the one-way Hello	• Helper does not receive the grace LSA from Restarter before the neighbor relationship expires. • Status of the interface that functions as the Helper changes. • Helper receives the LSA that

Execution of GR	Restarter	Helper
	packet from the Helper. • The Restarter receives the grace LSA that is generated by another router on the same network segment. Only one router can perform GR on the same network segment. • On the same network segment, neighbors of the Restarter have different DRs or BDRs because of the topology changes.	is inconsistent with the LSA in the local LSDB from another router. This situation can be excluded after the Helper is configured not to perform strict LSA check. • Helper receives grace LSAs from two routers on the same network segment at the same time. • Neighbor relationships between Helper and other neighbors change.

Comparison Between GR Mode and Non-GR Mode

Table 16-26 Comparison of master/slave switchover in the GR mode and non-GR mode

Switchover in Non-GR Mode	Switchover in GR Mode
• OSPF neighbor relationships are re-established. • Routes are recalculated. • Forwarding table changes. • Entire network detects route changes, and route flapping occurs for a short period of time. • Packets are lost during forwarding, and services are interrupted.	• OSPF neighbor relationships are re-established. • Routes are recalculated. • Forwarding table remains unchanged. • Except for neighbors of the device where master/slave switchover occurs, other routers do not detect route changes. • No packets are lost during forwarding, and services are not affected.

16.9.4 OSPF NSSA

Definition

OSPF Not-So-Stubby Areas (NSSA) are a new type of OSPF areas.

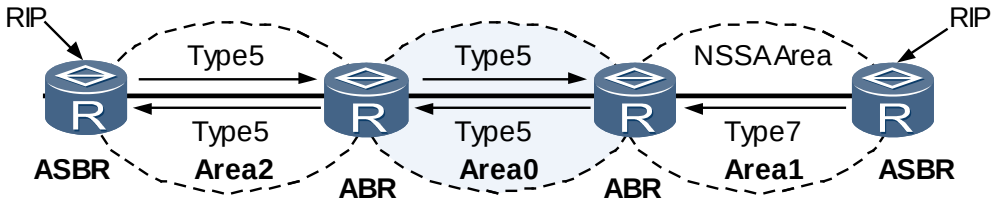
Derived from stub areas, NSSAs resemble stub areas in many ways. The difference between NSSAs and stub areas is that NSSAs can import and flood AS external routes to the entire OSPF AS, without learning external routes in other areas of the OSPF network.

Purpose

As defined in OSPF, stub areas cannot import external routes. This prevents a large number of external routes from consuming the bandwidth and storage resources of the Router s in stub

areas. Stub areas cannot meet the requirement of the scenario where external routes need to be imported Resource consumption caused by external routes also needs to be avoided. Therefore, NSSAs are introduced into the network.

Figure 16-52 NSSA



Type7 LSA

- Type7 LSAs are a new type of LSA that was introduced to support NSSAs and describe the imported external routes.
- Type7 LSAs are generated by the ASBRs of NSSAs and flooded only in the NSSAs where ASBRs reside.
- When receiving Type7 LSAs, the ABRs of NSSAs selectively translate Type7 LSAs to Type5 LSAs so that external routes can be advertised in other areas of the OSPF network.
- Default routes can also be expressed through Type7 LSAs so that traffic can be forwarded to other ASs.

N-bit

A Router in an area must be configured with the same area type. In OSPF, the N-bit is carried in a Hello packet to identify that a Router supports NSSAs. OSPF neighbor relationships cannot be established between the Routers with different area types.

Going against RFC 1587, some manufacturers also set the N-bit in OSPF Database Description (DD) packets. Huawei devices can be configured to be compatible with the devices of these manufacturers for interworking.

Translating Type7 LSAs to Type5 LSAs

To advertise the external routes imported by NSSAs in other areas, you need to translate Type7 LSAs to Type5 LSAs, so the external routes can be advertised in the entire OSPF network.

- The Propagate bit (P-bit) notifies a Router when Type7 LSAs need to be translated.
- The ABR with the largest Router ID in an NSSA translates Type7 LSAs to Type5 LSAs.
- Only the Type7 LSAs with the set P-bit and forwarding address other than 0 are translated to Type5 LSAs.

NOTE

FA indicates that the packet to a specific destination address is to be forwarded to the address specified by.

The loopback interface address in an area is preferentially selected as the FA. If no loopback interface exists, the address of the interface that is Up and has the largest logical index in the area is selected as the FA.

- Default Type7 LSAs that meet the preceding conditions can also be translated.
- Type7 LSAs generated by ABRs are not set with the P-bit.

Preventing Loops Caused by Default Routes

There may be multiple ABRs in an NSSA. To prevent routing loops, ABRs do not calculate the default routes advertised by the peer.

16.9.5 BFD for OSPF

Definition

Bidirectional Forwarding Detection (BFD) is a mechanism to detect communication faults between forwarding engines.

To be specific, BFD detects connectivity of a data protocol on the same path between two systems. The path can be a physical link, logical link, or tunnel.

In BFD for OSPF, a BFD session is associated with OSPF. The BFD session immediately detects a link fault and then notifies OSPF of the fault. This speeds up the OSPF's response to the change of the network topology.

Purpose

A link fault or topology change can cause Routers to recalculate routes. The convergence of routing protocols must be sped up to improve network performance.

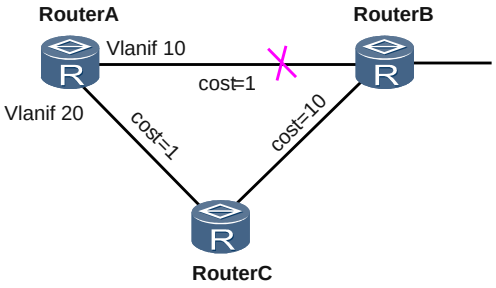
Because link faults are unavoidable, a feasible solution is required to detect faults faster and notify the faults to routing protocols immediately. If BFD is associated with routing protocols, when a link fault occurs, BFD can speed up the convergence of routing protocols.

Table 16-27 BFD for OSPF

Associated with BFD or Not	Link Fault Detection Mechanism	Convergence Speed
Not associated	An OSPF Dead timer expires. By default, the timeout period is 40s.	Seconds level
Associated	A BFD session goes Down.	At milliseconds level

Principle

Figure 16-53 BFD for OSPF



The principle of BFD for OSPF is shown in Figure 16-53.

1. OSPF neighbor relationships are established between the three Routers.
2. When a neighbor relationship becomes Full, this triggers BFD to establish a BFD session.
3. The outbound interface on Router A connected to Router B is GE 2. If the link fails, BFD detects the fault and notifies Router A.
4. Router A processes the event that a neighbor relationship is Down and re-calculates routes. After calculation, the outbound interface is GE 1 passes through Router and then reaches Router B.

16.9.6 OSPF Smart-discover

Definition

Generally, routers periodically send Hello packets through OSPF interfaces. That is, a router sends a Hello packet at the Hello interval set by a Hello timer. Sending Hello packets at a fixed interval slows down the establishment of OSPF neighbor relationships.

You can enable Smart-discover to speed up the establishment of OSPF neighbor relationships in specific scenarios.

Table 16-28 OSPF Smart-discover

Smart-discover Is Configured or Not	Processing
Not Configured	• Hello packets are sent only when the Hello timer expires. • Neighbors keep waiting to receive Hello packets within the timeout period.
Configured	• Hello packets are sent directly regardless of whether the Hello timer expires. • Neighbors can receive packets rapidly and perform status transition fast.

Principle

In the following scenarios, the interface enabled with Smart-discover can send Hello packets to neighbors actively, without having to wait for the Hello timer to expire:

- Neighbor status becomes two-way for the first time.
- Neighbor status changes from two-way or a higher state to Init.

16.9.7 OSPF-BGP Association

Definition

When a new router is deployed in the network or a router is restarted, network traffic may be lost during BGP convergence, because IGP convergence is faster than BGP convergence.

Establishing an association between OSPF and BGP can solve this problem.

Purpose

If a backup link exists, during traffic switchback, BGP traffic is lost because BGP route convergence is slower than OSPF route convergence.

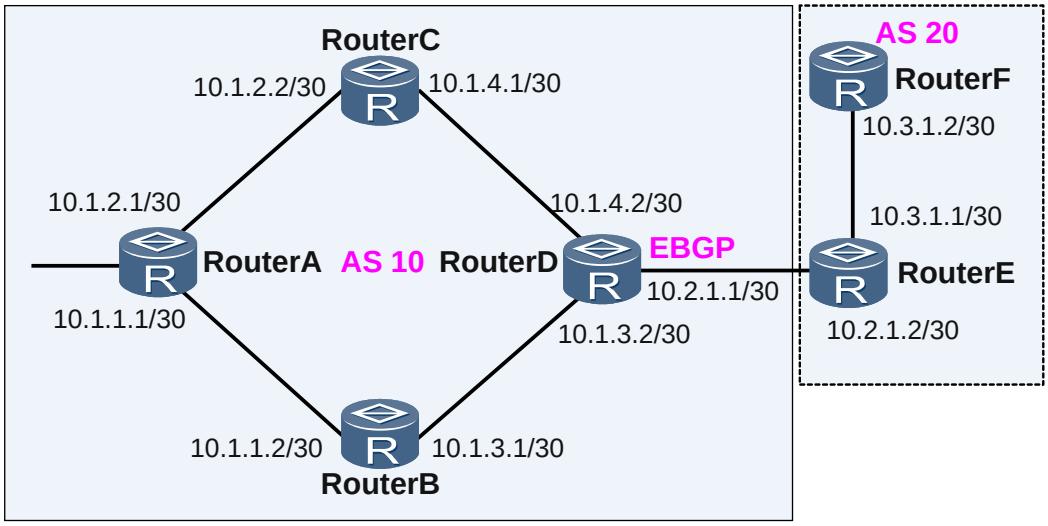
As shown in Figure 16-54, Router A, Router B, Router C, and Router D run OSPF and establish IBGP connections. Router C is the backup device of Router B. When the network is stable, BGP and OSPF routes converge completely on the devices.

Normally, traffic from Router A to address 10.3.1.0/30 passes through Router B. When Router B is faulty, traffic is switched to Router C. After Router B recovers, traffic is switched back to Router B, but packet loss occurs.

When traffic is switched back to Router B, IGP route convergence is again faster than BGP route convergence. OSPF routes converge first, whereas BGP route convergence is not complete. As a result, Router B has no way to reach 10.3.1.0/30.

When packets from Router A to 10.3.1.0/30 are sent to Router B, they are discarded by Router B because Router B has no route to 10.3.1.0/30.

Figure 16-54 OSPF-BGP association



Principle

In Figure 16-54, OSPF-BGP association is enabled on Router B. In this situation, before BGP route convergence is complete, Router A continues to forward traffic to the backup link Router C, without forwarding traffic to Router B, until BGP route convergence on Router B is complete.

The router enabled with OSPF-BGP association remains as a stub router within the set association period. That is, the link metric in the LSA advertised by the router is the maximum value of 65535. In this manner, the router notifies other OSPF devices not to use it as a transit router for data forwarding.

16.9.8 OSPF Database Overflow

Definition

OSPF requires that routers in the same area have the same Link State Database (LSDB).

With the continuous increase in routes on the network, some routers fail to carry the additional routing information because of limited system resources. This situation is called *OSPF database overflow*.

Purpose

You can configure stub areas or NSSAs to solve the problem of the continuous increase in routing information that causes the exhaustion of system resources of routers. However, configuring stub areas or NSSAs cannot solve the problem when the unexpected increase in dynamic routes causes database overflow. Setting the maximum number of external LSAs in the LSDB can dynamically limit the LSDB capacity, to avoid the problems caused by database overflow.

Principle

To prevent database overflow, you can set the maximum number of non-default external routes on a router.

All routers on the OSPF network must be set with the same upper limit. If the number of external routes on a router reaches the upper limit, the router enters the Overflow state and starts an overflow timer. The router automatically exits from the overflow state after the timer expires, By default, it is 5 seconds.

Table 16-29 OSPF database overflow

Overflow Phase	OSPF Processing
Entering overflow state	A router deletes all non-default external routes that are generated.
Staying in overflow state	• Router does not generate non-default external routes. • Router discards the newly received, non-default external routes, and does not reply with an LSAck packet. • When the overflow timer expires, the router checks whether the number of external routes still exceeds the upper limit. – If so, the router restarts the timer. – If not, the router exits from overflow state.
Exiting from the overflow state	• Router deletes the overflow timer. • Router generates non-default routes. • Router learns the newly received non-default routes, and replies with an LSAck packet. • Router prepares to enter Overflow state for the next time it occurs.

16.9.9 OSPF Fast Convergence

OSPF fast convergence is an extended feature of OSPF implemented to speed up the convergence of routes. It includes the following components:

- Incremental SPF (I-SPF): recalculates only the routes of the changed nodes rather than all the nodes when the network topology changes. This speeds up route calculation.
- Partial Route Calculation (PRC): calculates only the changed routes when the routes on the network change.
- An OSPF intelligent timer: can dynamically adjust its value based on the user's configuration and the interval at which an event is triggered, such as the route calculation interval, which ensures rapid and stable network operation.

  OSPF intelligent timer applies the exponential backoff technology so that the value of the timer can reach the millisecond level.

I-SPF

In ISO 10589, the Dijkstra algorithm was adopted to calculate routes. When a node changes on the network, this algorithm is used to recalculate all routes. The calculation takes a long time and consumes too many CPU resources, which affects the convergence speed.

I-SPF improves the Dijkstra algorithm. Except for the first time, only changed nodes instead of all nodes are involved in calculation. The SPT generated at last is the same as that generated by the previous algorithm, but I-SPF decreases CPU usage and speeds up network convergence.

PRC

Similar to I-SPF, PRC calculates only the changed routes. However, PRC does not calculate the shortest path. It updates routes based on the SPT calculated by I-SPF.

In route calculation, a leaf represents a route, and a node represents a router. SPT and leaf changes both cause the change of routing information, but the SPT change is irrelevant to the leaf change. PRC processes routing information based on either SPT or leaf information.

- If the SPT changes, PRC processes the routing information of all leaves on a changed node.
- If the SPT does not change, PRC does not process the routing information on any node.
- If a leaf changes, RPC processes the routing information on the leaf only.
- If a leaf does not change, PRC does not process the routing information on any leaf.

For example, if OSPF is enabled on an interface of a node, the SPT calculated by I-SPF remains unchanged. PRC updates only the routes of this interface, consuming less CPU resources.

PRC improves on the SPF algorithm. Working with I-SPF, RPC further improves the convergence performance of the network.

NOTE
In the implementation of a device, only I-SPF and PRC are used to calculate OSPF routes.

OSPF Intelligent Timer

On an unstable network, routes are calculated frequently, which consumes a great number of CPU resources. In addition, LSPs that describe the unstable topology are generated and transmitted on the unstable network. Frequently processing such LSAs affects the rapid and stable operation of the entire network.

To speed up route convergence on the entire network, the OSPF intelligent timer controls route calculation, LSA generation, and LSA receiving.

OSPF intelligent timer speeds up route convergence in the following modes:

- On a network where routes are calculated repeatedly, the OSPF intelligent timer dynamically adjusts the route calculation based on user's configuration and the exponential backoff technology. The number of route calculation times and the CPU resource consumption are decreased. Routes are calculated after the network topology becomes stable.
- On an unstable network, if a router generates or receives LSAs due to frequent topology changes, the OSPF intelligent timer can dynamically adjust its value. No LSA is

generated or handled within an interval, which prevents invalid LSAs from being generated and advertised on the entire network.

The OSPF intelligent timer is started by default and uses the default value.

16.9.10 OSPF Mesh-Group

Definition

In the scenario where there are multiple concurrent links, you can deploy OSPF mesh-group to classify links into a mesh group. Then, OSPF floods LSAs to only a link selected from the mesh group. Using OSPF mesh-group prevents unnecessary burden on the system caused by repetitive flooding.

The mesh-group feature is disabled by default.

Purpose

After receiving or generating an LSA, an OSPF process floods the LSA. When there are multiple concurrent links, OSPF floods the LSA to each link and sends Update messages.

In this scenario, if there are 2000 concurrent links, OSPF floods each LSA 2000 times. Only one flooding, however, is valid. The other 1999 times are useless repetition.

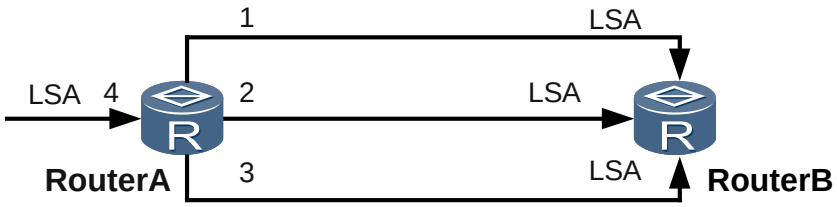
To prevent burden on the system caused by repetitive flooding, you can enable mesh-group to classify multiple concurrent links between a router and its neighbor into a group and then select a primary link to use for flooding.

Principles

As shown in Figure 16-55, Router A and Router B, which are connected through three links, establish an OSPF neighbor relationship. After receiving a new LSA from interface 4, Router A floods the LSA to Router B through interfaces 1, 2, and 3.

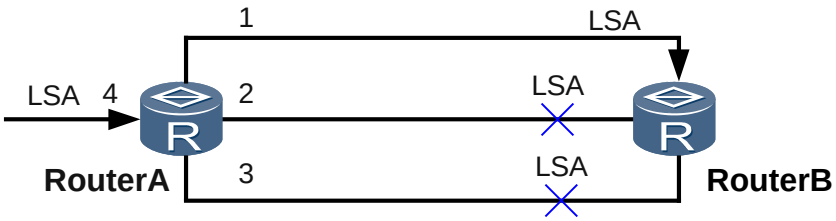
This flooding causes a heavy load on the concurrent links. For the neighbor with concurrent links, only a primary link is selected to flood the LSA.

Figure 16-55 LSAflooding with OSPF mesh-group disabled



When multiple concurrent links exist between a device enabled with OSPF mesh-group and its neighbor, the device selects? to flood the received LSAs, as shown in Figure 16-56.

Figure 16-56 LSA flooding with OSPF mesh-group enabled

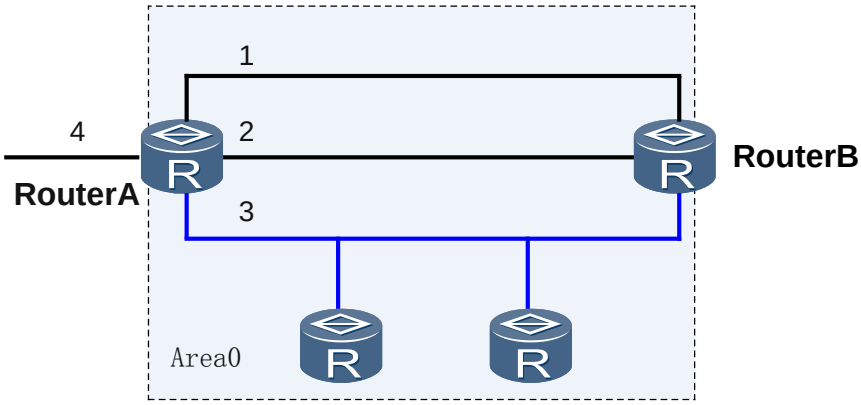


As defined in OSPF, LSAs can be flooded to a link only when the neighbor status is not lower than Exchange. In this case, when the status of the interface on the primary link is lower than Exchange, OSPF reselects a primary link from the concurrent links and then floods the LSA. After receiving the LSA flooded by Router A from link 1, Router B no longer floods the LSA to Router A through interfaces 2 and 3.

As defined by the mesh-group feature, the Router ID of a neighbor uniquely identifies the mesh group. Interfaces connected to the same neighbor that have a status greater than Exchange belong to the same mesh group.

In Figure 16-57, a mesh group of Router A resides in Area 0, which contains the links of interface 1 and interface 2. More than one neighbor of interface 3 resides on the broadcast link. Therefore, interface 3 cannot be defined as part of the mesh group.

Figure 16-57 Interface not added to mesh group



NOTE

After a router is enabled with mesh-group, if the Router IDs of the router and its directly connected neighbor are the same, LSDBs cannot be synchronized and routes cannot be calculated correctly. In this case, you need to reconfigure the Router ID of the neighbor.

16.9.11 Priority-based OSPF Convergence

Priority-based OSPF convergence ensures that specific routes converge first in the case of a great number of routes. Different routes can be set with different convergence priorities. This allows important routes to converge first and improves network reliability.

Using priority-based OSPF convergence, you can assign a high convergence priority to routes for key services so that those routes can converge quickly and minimize the impact on key services.

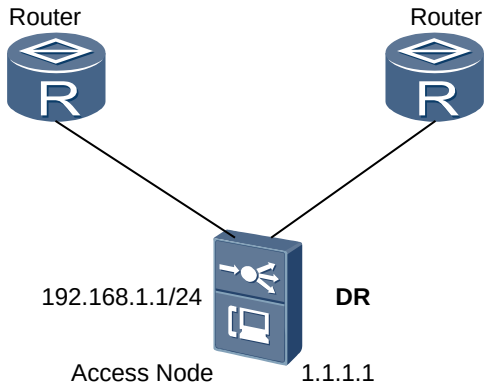
16.9.12 Configuration Example of OSPF

This topic provides an example for configuring OSPF on the MA5600T/MA5603T/MA5608T.

Service Requirements

- OSPF is enabled on the MA5600T/MA5603T/MA5608T.

Figure 16-58 Example network for configuring OSPF



Data Plan

Table 16-30 provides the data plan for configuring OSPF.

Table 16-30 Data plan for configuring OSPF

Item	Data
MA5600T/MA5603T/MA5608T	IP address of the Layer 3 interface: 192.168.1.1/24
	Priority: 100
	VLAN ID: 2
	Router ID: 1.1.1.1

Context

- The native VLAN of each interface of the MA5600T/MA5603T/MA5608T must be configured to ensure a normal communication.
- The OSPF area IDs of the MA5600T/MA5603T/MA5608T device and the routers must be consistent.

Procedure

Configure MA5600T/MA5603T/MA5608T.

1. Configure the IP address of the Layer 3 interface.

```
huawei(config)#vlan 2 smart
huawei(config)#port vlan 2 0/19 0
huawei(config)#interface vlanif 2
huawei(config-if-vlanif2)#ip address 192.168.1.1 24
huawei(config-if-vlanif2)#quit
```

2. Configure the OSPF Router ID.

```
huawei(config)#router id 1.1.1.1
```

3. Enable OSPF.

```
huawei(config)#ospf
huawei(config-ospf-1)#area 0
huawei(config-ospf-1-area-0.0.0.0)#network 192.168.1.0 0.0.0.255
huawei(config-ospf-1-area-0.0.0.0)#network 1.1.1.1 0.0.0.0
huawei(config-ospf-1-area-0.0.0.0)#quit
huawei(config-ospf-1)#quit
```

4. Configure the OSPF priority.

```
huawei(config)#interface vlanif 2
huawei(config-if-vlanif2)#ospf dr-priority 100
huawei(config-if-vlanif2)#quit
```

5. Save the data.

```
huawei(config)#save
```

----End

Result

Run the **display ip routing-table** command and you can find the learnt route table. Hosts can communicate with each other.

Configuration File

```
vlan 2 smart
port vlan 2 0/19 0
interface vlanif 2
ip address 192.168.1.1 24
quit
router id 1.1.1.1
ospf
area 0
network 192.168.1.0 0.0.0.255
network 1.1.1.1 0.0.0.0
quit
quit
interface vlanif 2
ospf dr-priority 100
quit
save
```

16.9.13 References

The following table lists the references that apply in this chapter.

Document No.	Document Name	Protocol Compliance
RFC 1587	The OSPF NSSA Option	Fully compliant.
RFC 1765	OSPF Database Overflow	Fully compliant. This RFC is experimental and non-standard.
RFC 2328	OSPF Version 2	Fully compliant.
RFC 2370	The OSPF Opaque LSA Option	Fully compliant.
RFC 3137	OSPF Stub Router Advertisement	Fully compliant. This RFC is informational and non-standard.
RFC 3623	Graceful OSPF Restart	Fully compliant.
RFC 3630	Traffic Engineering (TE) Extensions to OSPF Version 2	Fully compliant.
RFC 3682	The Generalized TTL Security Mechanism (GTSM)	Fully compliant. This RFC is experimental and non-standard.
RFC 3906	Calculating Interior Gateway Protocol (IGP) Routes Over Traffic Engineering Tunnels	Fully compliant.
RFC 4576	Using a Link State Advertisement (LSA) Options Bit to Prevent Looping in BGP/MPLS IP Virtual Private Networks (VPNs)	Fully compliant.
RFC 4577	OSPF as the Provider/Customer Edge Protocol for BGP/MPLS IP Virtual Private Networks (VPNs)	Fully compliant.
RFC 4750	OSPF Version 2 Management Information Base	Partially compliant. The SET operation is not supported.

16.10 OSPFv3

16.10.1 Introduction to OSPFv3

Definition

The Open Shortest Path First (OSPF) protocol, developed by the Internet Engineering Task Force (IETF), is an interior gateway protocol based on the link status.

At present, OSPF Version 2 is used for IPv4 and OSPF Version 3 is used for IPv6.

- OSPFv3 is short for OSPF Version 3.
- As defined in RFC 2740, OSPFv3 is a routing protocol over IPv6.
- OSPFv3 is an independent routing protocol whose functions are enhanced on the basis of OSPFv2.

Purpose

The primary purpose of OSPFv3 is to develop a routing protocol independent of any specific network layer. The internal routing information of OSPFv3 is redesigned to serve this purpose.

The differences between OSPFv3 and OSPFv2 are as follows:

- OSPFv3 does not insert IP-based data in the header of the packet and Link State Advertisement (LSA).
- OSPFv3 executes some crucial tasks that originally require the data in the IP packet header by making use of the information independent of any network protocol. For example, OSPFv3 can identify the LSA that advertises the routing data.

16.10.2 Principle of OSPFv3

Running on IPv6, OSPFv3 (defined in RFC 2740) is an independent routing protocol whose functions are enhanced on the basis of OSPFv2.

- OSPFv3 and OSPFv2 are the same in respect of the working principles of the Hello message, state machine, link-state database (LSDB), flooding, and route calculation.
- OSPFv3 divides an Autonomous System (AS) into one or more logical areas and advertises routes through LSAs.
- OSPFv3 achieves unity of routing information by exchanging OSPFv3 packets between routers within an OSPFv3 area.
- OSPFv3 packets are encapsulated into IPv6 packets, which can be transmitted in unicast or multicast mode.

Formats of OSPFv3 Packets

Packet Type	Description
Hello message	Hello messages are sent regularly to discover and maintain OSPFv3 neighbor relationships.
Database Description (DD) packet	A DD packet contains the summary of the local LSDB. It is exchanged between two OSPFv3 routers to update the LSDBs.
Link State Request (LSR) packet	LSR packets are sent to the neighbor to request the required LSAs. An OSPFv3 router sends LSR packets to its neighbor only after they exchange DD packets.
Link State Update (LSU) packet	The LSU packet is used to transmit required LSAs to the neighbor.
Link State Acknowledgment	The LSAck packet is used to acknowledge the received

Packet Type	Description
(LSAck) packet	LSA packets.

LSA Type

LSA Type	Description
Router-LSA (Type1)	Generated by a router for each area to which an OSPFv3 interface belongs, the router LSA describes the status and costs of links of the router and is advertised in the area where the OSPFv3 interface belongs.
Network-LSA (Type2)	Generated by a designated router (DR), the network LSA describes the link status and is broadcast in the area that the DR belongs to.
Inter-Area-Prefix-LSA (Type3)	Generated on the area border router (ABR), an inter-area prefix LSA describes the route of a certain network segment within the local area and is used to inform other areas of the route.
Inter-Area-Router-LSA (Type4)	Generated on the ABR, an inter-area router LSA describes the route to the autonomous system boundary router (ASBR) and is advertised to all related areas except the area that the ASBR belongs to.
AS-external-LSA (Type5)	Generated on the ASBR, the AS-external LSA describes the route to a destination outside the AS and is advertised to all areas except the stub area and NSSA area.
NSSA-LSA (Type7)	Describes routes to a destination outside the AS. It is generated by an ASBR and advertised in NSSAs only.
Link-LSA (Type8)	Each router generates a link LSA for each link. A link LSA describes the link-local address and IPv6 address prefix associated with the link and the link option set in the network LSA. It is transmitted only on the link.
Intra-Area-Prefix-LSA (Type9)	Each router or DR generates one or more intra-area prefix LSAs and transmits it in the local area. • An LSA generated on a router describes the IPv6 address prefix associated with the router LSA. • An LSA generated on a DR describes the IPv6 address prefix associated with the network LSA.

Router Type

Figure 16-59 Router type

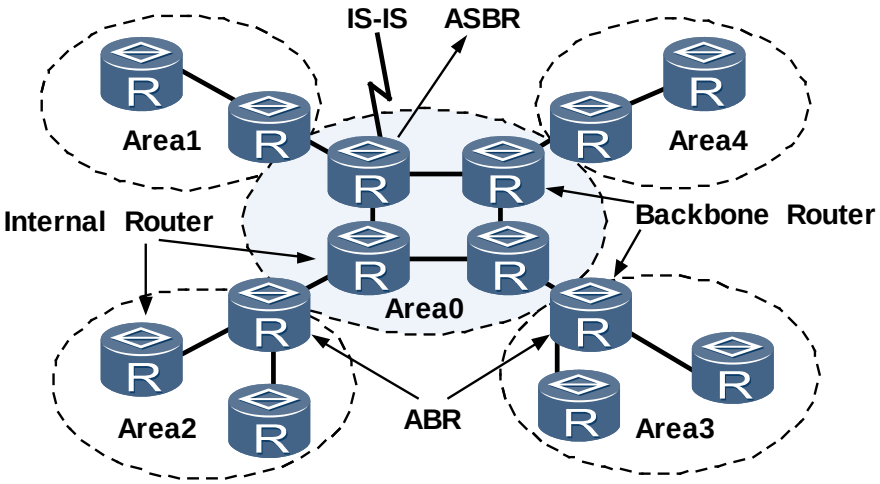


Table 16-31 Router types and descriptions

Router Type	Description
Internal router	All interfaces on an internal router belong to the same OSPFv3 area.
Area border router (ABR)	An ABR can belong to two or more areas, but one of the areas must be a backbone area. An ABR is used to connect the backbone area and the non-backbone areas. It can be physically or logically connected to the backbone area.
Backbone router	At least one interface on a backbone router belongs to the backbone area. All ABRs and internal routers in Area 0, therefore, are backbone routers.
AS boundary router (ASBR)	A router that exchanges routing information with other ASs is called an ASBR. An ASBR may not locate on the boundary of an AS. It can be an internal router or an ABR.

OSPFv3 Route Type

Inter-area routes and intra-area routes describe the network structure of an AS. External routes describe how to select a route to the destination outside an AS. OSPFv3 classifies the imported AS external routes into Type 1 routes and Type 2 routes.

Table 16-32 lists route types in a descending order of priority.

Table 16-32 Types of OSPFv3 routes

Route Type	Description
Intra Area	Intra-area routes
Inter Area	Inter-area routes
Type1 external routes	Because of the high reliability of Type 1 external routes, the calculated cost of external routes is equal to that of AS internal routes, and can be compared with the cost of OSPFv3 routes. That is, the cost of a Type1 external route equals the cost of the route from the router to the corresponding ASBR plus the cost of the route from the ASBR to the destination address.
Type2 external routes	Because of the low reliability of Type2 external routes, the cost of the route from the ASBR to a destination outside the AS is considered far greater than the cost of any internal path to an ASBR. Therefore, OSPFv3 only takes the cost of the route from the ASBR to a destination outside the AS into account when calculating route costs. That is, the cost of a Type2 external route equals the cost of the route from the ASBR to the destination of the route.

Area

When a large number of Routers run OSPFv3, link state databases (LSDBs) become very large and require a large amount of storage space. Large LSDBs also complicate shortest path first (SPF) computation and are computationally intensive for the Routers. Network expansion causes the network topology to change, which results in route flapping and frequent OSPFv3 packet transmission. When a large number of OSPFv3 packets are transmitted on the network, bandwidth usage efficiency decreases. Each change in the network topology causes all Routers on the network to recalculate routes.

OSPFv3 resolves this problem by partitioning an AS into different areas. An area is regarded as a logical group, and each group is identified by an area ID. A Router, not a link, resides at the border of an area. A network segment or link can belong only to one area. An area must be specified for each OSPFv3 interface.

OSPFv3 areas include common areas, stub areas, and not-so-stubby areas (NSSAs), as described in Table 16-33.

Table 16-33 OSPF areas

Area Type	Function	Notes
Common area	By default, OSPFv3 areas are defined as common areas. Common areas include: • Standard area: transmits intra-area, inter-area, and external routes.	• The backbone area must have all its devices connected. • All non-backbone

Area Type	Function	Notes
	• Backbone area: connects to all other OSPFv3 areas and transmits inter-area routes. The backbone area is represented by area 0. Routes between non-backbone areas must be forwarded through the backbone area.	areas must remain connected to the backbone area.
Stub area	A stub area is a non-backbone area with only one ABR and generally resides at the border of an AS. The area border router (ABR) in a stub area does not transmit received AS external routes, which significantly decreases the number of entries in the routing table on the ABR and the amount of routing information to be transmitted. To ensure the reachability of AS external routes, the ABR in the stub area generates a default route and advertises the route to non-ABRs in the stub area. A totally stub area allows only intra-area routes and ABR-advertised Type 3 link state advertisements (LSAs) carrying a default route to be advertised within the area.	• The backbone area cannot be configured as a stub area. • An autonomous system boundary router (ASBR) cannot exist in a stub area. Therefore, AS external routes cannot be advertised within the stub area. • A virtual link cannot pass through a stub area.
NSSA	An NSSA is similar to a stub area. An NSSA does not advertise Type 5 LSAs but can import AS external routes. ASBRs in an NSSA generate Type 7 LSAs to carry the information about the AS external routes. The Type 7 LSAs are advertised only within the NSSA. When the Type 7 LSAs reach an ABR in the NSSA, the ABR translates the Type 7 LSAs into Type 5 LSAs and floods them to the entire AS. A totally NSSA area allows only intra-area routes to be advertised within the area.	• ABRs in an NSSA advertise Type 3 LSAs carrying a default route within the NSSA. All inter-area routes are advertised by ABRs. • A virtual link cannot pass through an NSSA.

Network Types Supported by OSPFv3

OSPFv3 classifies networks into the following types according to link layer protocols.

Table 16-34 Types of OSPFv3 networks

Network Type	Description
Broadcast	If the link layer protocol is Ethernet or FDDI, OSPFv3 defaults the network type to broadcast. In this type of networks, the following situations occur: • Hello messages, LSU packets, and LSAck packets are transmitted in multicast mode (FF02::5 is the reserved IPv6 multicast address of the OSPFv3 router; FF02::6 is the reserved IPv6 multicast address of the OSPFv3 DR or BDR).

Network Type	Description
	• DD packets and LSR packets are transmitted in unicast mode.
Non-broadcast Multiple Access (NBMA)	If the link layer protocol is frame relay, ATM, or X.25, OSPFv3 defaults the network type to NBMA. In this type of networks, protocol packets such as Hello messages, DD packets, LSR packets, LSU packets, and LSAck packets, are transmitted in unicast mode.
Point-to-Multipoint (P2MP)	Regardless of the link layer protocol, OSPFv3 does not default the network type to P2MP. A P2MP network must be forcibly changed from other network types. The common practice is to change a non-fully connected NBMA to a P2MP network. In this type of networks, the following situations occur: • Hello messages are transmitted in multicast mode with the multicast address as FF02::5. • Other protocol packets, including DD packets, LSR packets, LSU packets, and LSAck packets, are transmitted in unicast mode.
Point-to-point (P2P)	If the link layer protocol is PPP, HDLC, or LAPB, OSPFv3 defaults the network type to P2P. In this type of network, the protocol packets, including Hello messages, DD packets, LSR packets, LSU packets, and LSAck packets, are transmitted to the multicast address FF02::5.

Stub Area

A stub area is a special area where the ABRs do not flood the received external routes. In stub areas, the size of the routing table of the routers and the routing information in transmission are reduced.

Configuring a stub area is optional. Not all areas can be configured as stub areas. Usually, a stub area is a non-backbone area with only one ABR and is located at the AS boundary.

To ensure the reachability of a destination outside the AS, the ABR in the stub area generates a default route and advertises it to the non-ABR routers in the stub area.

Note the following when configuring a stub area:

- The backbone area cannot be configured as a stub area.
- If an area needs to be configured as a stub area, all the routers in this area must be configured with the **stub** command.
- An ASBR cannot exist in a stub area. That is, external routes are not flooded in the stub area.
- A virtual link cannot pass through the stub area.

OSPFv3 Route Summarization

Routing information can be decreased after route aggregation so that the size of routing tables is reduced, which improves the performance of routers.

The procedure for OSPFv3 route aggregation is as follows:

- Route summarization on an ABR

  An ABR can summarize routes with the same prefix into one route and advertise the summarized route in other areas.

  When sending routing information to other areas, an ABR generates Type 3 LSAs based on IPv6 prefixes. If consecutive IPv6 prefixes exist in an area and route summarization is enabled on the ABR of the area, the IPv6 prefixes can be summarized into one prefix. If there are multiple LSAs that have the same prefix, the ABR summarizes these LSAs and advertises only one summarized LSA. The ABR does not advertise any specific LSAs.
- Route summarization on an ASBR

  An ASBR can summarize imported routes with the same prefix into one route and then advertise the summarized route to other areas.

  After being enabled with route summarization, an ASBR summarizes imported Type 5 LSAs within the summarized address range. After route summarization, the ASBR does not generate a separate Type 5 LSA for each specific prefix within the configured range. Instead, the ASBR generates a Type 5 LSA for only the summarized prefix. In an NSSA, an ASBR summarizes multiple imported Type 7 LSAs within the summarized address range into one Type 7 LSA.

OSPFv3 Virtual Link

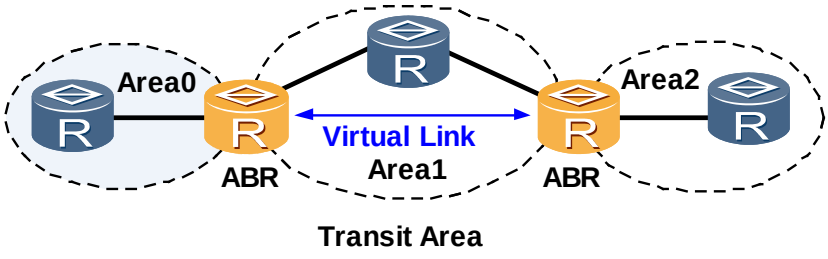
A virtual link refers to a logical channel established between two ABRs through a non-backbone area.

- A virtual link must be set up on both ends of the link; otherwise, it does not take effect.
- The transmit area refers to the area that provides an internal route of a non-backbone area for both the ends of the virtual link.

In actual applications, the physical connectivity between non-backbone areas and the backbone area cannot be ensured owing to various limitations. To solve this problem, you can configure OSPFv3 virtual links.

The virtual link is similar to a point-to-point connection between two ABRs. Similar to physical interfaces, the interfaces on the virtual link can be configured with parameters such as the hello interval.

Figure 16-60 OSPFv3 virtual link



As shown in Figure 16-60, OSPFv3 packets transmitted between two ABRs are only forwarded by the OSPFv3 devices that reside between the two ABRs. The OSPFv3 devices detect that they are not the destinations of the packets, so they forward the packets as common IP packets.

OSPFv3 Multi-process

OSPFv3 supports multi-process. More than one OSPFv3 process can run on the same router because processes are independent of each other. Route interaction between different OSPFv3 processes is similar to the route interaction between different routing protocols.

An interface of a router belongs to only a certain OSPFv3 process.

16.10.3 OSPFv3 GR

Graceful restart (GR) is a technology used to ensure normal traffic forwarding when a routing protocol restarts and guarantee that key services are not affected in the process.

GR is one of the high availability (HA) technologies, which comprise a series of comprehensive technologies such as fault-tolerant redundancy, link protection, faulty node recovery, and traffic engineering. As a redundancy technology, GR is widely used to ensure uninterrupted forwarding of key data in active/standby switchover and system upgrade.

- If GR is not enabled, the active/standby switchover occurring owing to various causes leads to transient interruption of data forwarding, and as a result, route flapping occurs on the whole network. Such route flapping and service interruption are unacceptable on a large-scale network, especially on a carrier network.

In GR mode, the forwarding plane continues to direct data forwarding once a restart occurs, and the actions on the control plane, such as reestablishment of neighbor relationships and route calculation, do not affect the forwarding plane. In this manner, service interruption caused by route flapping is prevented so that the network reliability is improved.

Basic Concepts

- Grace LSA
  - OSPFv3 supports GR by flooding grace LSAs on the link.
  - Grace LSAs are used to inform the neighbor of the GR time, cause, and interface instance ID when GR starts and ends.
- Router function
  - A router can function as a GR restarter.
  - A router can function as a GR helper.
- GR implementation
  - Planned-GR: This refers to the smooth restart of OSPFv3 through the **reset ospfv3 graceful-restart** command. In this mode, a grace LSA is sent to the neighbor before the restart.
  - Unplanned-GR: This refers to the active/standby switchover triggered by router faults like power down, dead loop, exception or reset in master.

    Unlike planned-GR, no grace LSA is sent before the active/standby switchover in unplanned GR mode. Instead, the switchover is directly performed. When the standby board becomes Up, a grace LSA is sent and the GR process starts. The following procedure is the same as that of planned GR.

GR Process

Figure 16-61 OSPFv3 planned-GR process (reset ospfv3 graceful-restart)

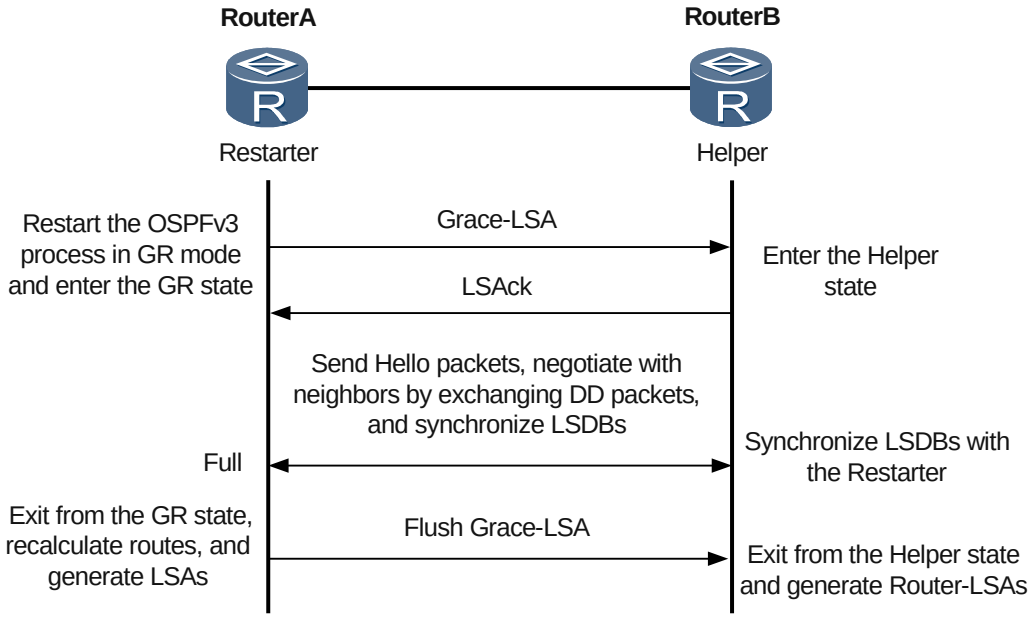
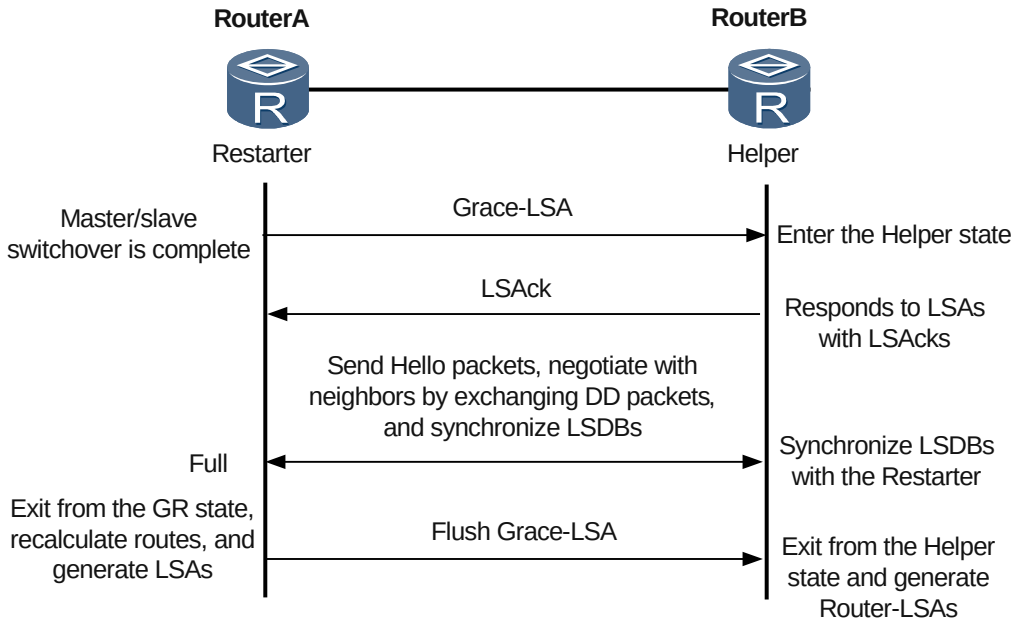


Figure 16-62 OSPFv3 unplanned-GR process (active/standby switchover)



- On the GR restarter:

1. In planned-GR mode, when OSPFv3 is restarted through commands, the GR restarter sends a grace LSA to all neighbors to inform them of the start of a GR process and the period and cause of this process.

In unplanned GR mode, when a restart occurs after active/standby switchover due to the unplanned causes (software failure, power failure) other than commands, a grace LSA is sent to each neighbor immediately after the standby board is Up to inform the neighbors of the start of a GR process and the period and cause of the process.

2. The GR restarter performs negotiation with neighbors again to set up new neighbor relationships.
3. When all the neighbor relationships between the GR restarter and the original neighbors enter the Full state:
   - The GR restarter exits from the GR process and OSPFv3 recalculates routes.
   - The GR restarter updates the routing table on the main control board and the FIBs on interface boards and deletes invalid routing entries.
   - The GR restarter sends a grace LSA whose aging time is 3600 seconds to instruct the GR helper to exit from the GR process.

   Now, the GR process is complete.
4. If errors occur, the GR timer expires, or the neighbor relationship fails to enter the Full state during a GR process, the GR restarter exits from the process and OSPFv3 is restarted in non-GR mode. In this case, packets are lost.

- On the GR helper:

1. If a router is configured to support the GR process on its neighbor, the router enters the helper mode after receiving a grace LSA.
2. The GR helper maintains its neighbor relationship with the GR restarter, and the status of the neighbor relationship does not change.
3. If the GR helper continues to receive grace LSAs whose GR period is different from that on the GR helper, the GR helper updates its GR period.
4. Being informed of the successful GR process through a grace LSA whose aging time is 3600 seconds from the GR restarter, the GR helper exits from the GR process.
5. If errors occur during a GR process, the GR helper exits from the helper state and deletes invalid routes after route calculation.

Comparison between the GR Mode and the Non-GR Mode

Table 16-35 Comparison between the GR mode and the non-GR mode

Active/Standby Switchover in Non-GR Mode	Active/Standby Switchover in GR Mode
• OSPFv3 neighbor relationships are reestablished. • Routes are recalculated. • The forwarding table changes. • Route changes are sensed on the network and route flapping occurs over a short period of time. • Packets are lost during forwarding, and services are interrupted.	• OSPFv3 neighbor relationships are reestablished. • Routes are recalculated. • The forwarding table remains the same. • Except the neighbor of the device where the active/standby switchover occurs, other routers do not sense the route changes. • No packets are lost during forwarding, and services are not affected.

16.10.4 BFD for OSPFv3

Definition

Bidirectional Forwarding Detection (BFD) is a mechanism used to detect faults of communications between forwarding engines.

To be specific, BFD detects connectivity of a data protocol on a path between two systems. The path can be a physical link, a logical link, or a tunnel.

BFD for OSPFv3 associates BFD with OSPFv3. BFD fast detects a link fault and then notifies OSPFv3 of the fault. This speeds up OSPFv3's response to the change of the network topology.

Purpose

A link fault or the topology change causes Routers to recalculate routes. Therefore, the convergence of routing protocols must be as quick as possible to improve network performance.

Link faults are inevitable. Therefore, fast detecting faults and notifying routing protocols of the faults is a feasible solution to immediately rectify link faults. After BFD is associated with routing protocols, BFD can speed up the convergence of routing protocols if a link fault occurs.

Principles

Figure 16-63 BFD for OSPFv3

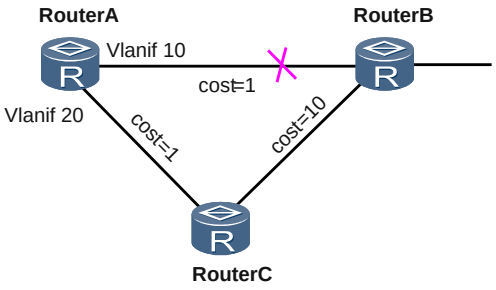


Figure 16-63 shows the principle of BFD for OSPFv3.

1. OSPFv3 neighbor relationships are established between these three Routers.
2. After a neighbor relationship becomes Full, this triggers BFD to establish a BFD session.
3. The outbound interface on Router A connected to Router B is Vlanif10. If the link fails, BFD detects the fault and then notifies Router A of the fault.
4. Router A processes the event that a neighbor relationship becomes Down and re-calculates routes. After calculation, the outbound interface is Vlanif20 passes through Router C and then reaches Router B.

16.10.5 Comparison between OSPFv3 and OSPFv2

OSPFv3 and OSPFv2 are the same in the following aspects:

- Network type and interface type
- Interface state machine and neighbor state machine
- LSDB
- Flooding mechanism
- Five types of packets, including Hello, DD, LSR, LSU, and LSAck packets
- Route calculation

OSPFv3 and OSPFv2 are different in the following aspects:

- OSPFv3 is based on links rather than network segments.

  OSPFv3 runs on IPv6, which is based on links rather than network segments.

  Therefore, you need not to configure OSPFv3 on the interfaces in the same network segment. It is only required that the interfaces enabled with OSPFv3 are on the same link. In addition, the interfaces can set up OSPFv3 sessions without IPv6 global addresses.
- OSPFv3 does not depend on IP addresses.

  This is to separate topology calculation from IP addresses. That is, OSPFv3 can calculate the OSPFv3 topology without knowing the IPv6 global address, which only applies to virtual link interfaces for packet forwarding.
- OSPFv3 packets and LSA format change.
  - OSPFv3 packets do not contain IP addresses.
  - OSPFv3 router LSAs and network LSAs do not contain IP addresses, which are advertised by link LSAs and intra-area prefix LSAs.
  - In OSPFv3, Router IDs, area IDs, and LSA link state IDs no longer indicate IP addresses, but the IPv4 address format is still reserved.
  - Neighbors are identified by Router IDs instead of IP addresses in broadcast, NBMA, or P2MP networks.
- Information about the flooding scope is added in LSAs of OSPFv3.

  Information about the flooding scope is added in the LSA Type field of LSAs of OSPFv3. Thus, OSPFv3 routers can process LSAs of unidentified types, which makes the processing more flexible.
  - OSPFv3 can store or flood unidentified packets, whereas OSPFv2 just discards unidentified packets.
  - OSPFv3 floods packets in an OSPF area or on a link. It sets the U flag bit of packets (the flooding area is based on the link local) so that unidentified packets are stored or forwarded to the stub area.

  For example, Router A and Router B can identify LSAs of a certain type. They are connected through Router C, which, however, cannot identify this type of LSAs. When Router A floods an LSA of this type, Router C can still flood the received LSA to Router B although it does not identify this LSA. Router B then processes the LSA.

  If OSPFv2 is run, Router C discards the unidentified LSA so that the LSA cannot reach Router B.
- OSPFv3 supports multi-process on a link.

  Only one OSPF process can be configured on a physical interface.

  In OSPFv3, one physical interface can be configured with multiple processes that are identified by different instance IDs. That is, multiple OSPFv3 instances can run on one physical link. They establish neighbor relationships with the other end of the link and transmit packets to the other end without interfering with each other.

Thus, the resources of a link can be shared among OSPFv3 instances that simulate multiple OSPFv3 routers, which improves the utilization of limited router resources.

- OSPFv3 uses IPv6 link-local addresses.

  IPv6 implements neighbor discovery and automatic configuration based on link-local addresses. Routers running IPv6 do not forward IPv6 packets whose destination address is a link-local address. Those packets can only be exchanged on the same link. The unicast link-local address starts from FE80/10.

  As a routing protocol running on IPv6, OSPFv3 also uses link-local addresses to maintain neighbor relationships and update LSDBs. Except Vlink interfaces, all OSPFv3 interfaces use link-local addresses as the source address and that of the next hop to transmit OSPFv3 packets.

  The advantages are as follows:

  - The OSPFv3 can calculate the topology without knowing the global IPv6 addresses so that topology calculation is not based on IP addresses.
  - The packets flooded on a link are not transmitted to other links, which prevents unnecessary flooding and saves bandwidth.

- OSPFv3 supports two new LSAs.
  - Link LSA: A router floods a link LSA on the link where it resides to advertise its link-local address and the configured global IPv6 address.
  - Intra-area prefix LSA: A router advertises an intra-area prefix LSA in the local OSPF area to inform the other routers in the area or the network, which can be a broadcast network or a NBMA network, of its IPv6 global address.

- OSPFv3 identifies neighbors based on router IDs only.

  On broadcast, NBMA, and P2MP networks, OSPFv2 identifies neighbors based on IPv4 addresses of interfaces.

  OSPFv3 identifies neighbors based on router IDs only. Thus, even if global IPv6 addresses are not configured or they are configured in different network segments, OSPFv3 can still establish and maintain neighbor relationships so that topology calculation is not based on IP addresses.

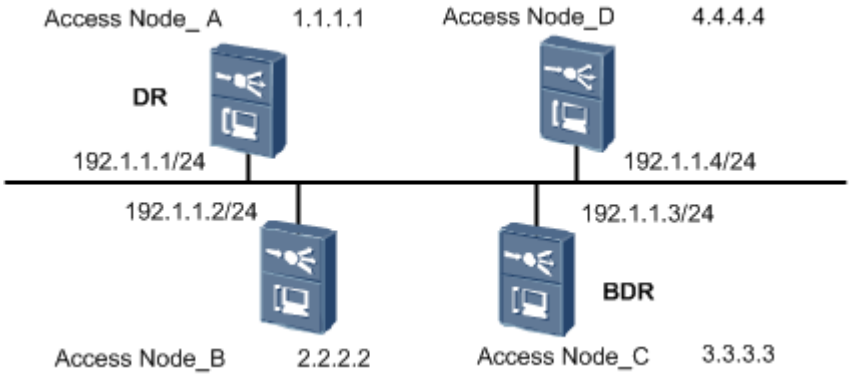
16.10.6 Configuration Example of OSPFv3

This topic provides an example for configuring OSPFv3 on the MA5600T/MA5603T/MA5608T.

Service Requirements

- OSPFv3 is enabled on the four MA5600T/MA5603T/MA5608Ts.
- Access Node_A is configured with the highest designated router (DR) priority, Access Node_C is configured with the second highest DR priority, and Access Node_A implements the broadcast of network link status for the DR.

Figure 16-64 Example network for configuring OSPFv3



Data Plan

Table 16-36 provides the data plan for configuring OSPFv3.

Table 16-36 Data plan for configuring OSPFv3

Item	Data	Remarks
Access Node_A	IPv6 address of the Layer 3 interface: 192:1::1:1/64	-
	Priority: 100	Ensure one of the access nodes is configured with the highest designated router (DR) priority. Take Access Node_A for example.
	VLAN ID: 2	-
	Router ID: 1.1.1.1	-
Access Node_B	IPv6 address of the Layer 3 interface: 192:1::1:2/64	-
	Priority: 80	-
	VLAN ID: 2	-
	Router ID: 2.2.2.2	-
Access Node_C	IPv6 address of the Layer 3 interface: 192:1::1:3/64	-
	Priority: 90	-
	VLAN ID: 2	-
	Router ID: 3.3.3.3	-
Access Node_D	IPv6 address of the Layer 3	-

Item	Data	Remarks
	interface: 192:1::1:4/64	
	Priority: not configured	Default: 1
	VLAN ID: 2	-
	Router ID: 4.4.4.4	-

Context

The OSPFv3 area IDs of the MA5600T/MA5603T/MA5608T devices must be consistent. Take area 0 for example.

Procedure

Configure Access Node_A.

1. Configure the IPv6 address of the Layer 3 interface.

```
huawei(config)#ipv6
huawei(config)#vlan 2
huawei(config)#port vlan 2 0/19 0
huawei(config)#interface vlanif 2
huawei(config-if-vlanif2)#ipv6 enable
huawei(config-if-vlanif2)#ipv6 address 192:1::1:1 64
huawei(config-if-vlanif2)#quit
```

2. Enable OSPFv3, and configure the OSPFv3 Router ID.

```
huawei(config)#ospfv3
huawei(config-ospfv3-1)#router-id 1.1.1.1
huawei(config-ospfv3-1)#quit
```

3. Enable OSPFv3 on the vlanif interface, and configure the OSPFv3 priority.

```
huawei(config)#interface vlanif 2
huawei(config-if-vlanif2)#ospfv3 1 area 0
huawei(config-if-vlanif2)#ospfv3 dr-priority 100
huawei(config-if-vlanif2)#quit
```

4. Save the data.

```
huawei(config)#save
```

Step 1 Configure Access Node_B.

1. Configure the IPv6 address of the Layer 3 interface.

```
huawei(config)#ipv6
huawei(config)#vlan 2
huawei(config)#port vlan 2 0/19 0
huawei(config)#interface vlanif 2
huawei(config-if-vlanif2)#ipv6 enable
huawei(config-if-vlanif2)#ipv6 address 192:1::1:2 64
huawei(config-if-vlanif2)#quit
```

2. Enable OSPFv3, and configure the OSPFv3 Router ID.

```
huawei(config)#ospfv3
huawei(config-ospfv3-1)#router-id 2.2.2.2
huawei(config-ospfv3-1)#quit
```

3. Enable OSPFv3 on the vlanif interface, and configure the OSPFv3 priority.

```
huawei(config)#interface vlanif 2
huawei(config-if-vlanif2)#ospfv3 1 area 0
huawei(config-if-vlanif2)#ospfv3 dr-priority 80
huawei(config-if-vlanif2)#quit
```

4. Save the data.

```
huawei(config)#save
```

Step 2 Configure Access Node_C.

1. Configure the IPv6 address of the Layer 3 interface.

```
huawei(config)#ipv6
huawei(config)#vlan 2
huawei(config)#port vlan 2 0/19 0
huawei(config)#interface vlanif 2
huawei(config-if-vlanif2)#ipv6 enable
huawei(config-if-vlanif2)#ipv6 address 192:1::1:3 64
huawei(config-if-vlanif2)#quit
```

2. Enable OSPFv3, and configure the OSPFv3 Router ID.

```
huawei(config)#ospfv3
huawei(config-ospfv3-1)#router-id 3.3.3.3
huawei(config-ospfv3-1)#quit
```

3. Enable OSPFv3 on the vlanif interface, and configure the OSPFv3 priority.

```
huawei(config)#interface vlanif 2
huawei(config-if-vlanif2)#ospfv3 1 area 0
huawei(config-if-vlanif2)#ospfv3 dr-priority 90
huawei(config-if-vlanif2)#quit
```

4. Save the data.

```
huawei(config)#save
```

Step 3 Configure Access Node_D.

1. Configure the IPv6 address of the Layer 3 interface.

```
huawei(config)#ipv6
huawei(config)#vlan 2
huawei(config)#port vlan 2 0/19 0
huawei(config)#interface vlanif 2
huawei(config-if-vlanif2)#ipv6 enable
huawei(config-if-vlanif2)#ipv6 address 192:1::1:4 64
huawei(config-if-vlanif2)#quit
```

2. Enable OSPFv3, and configure the OSPFv3 Router ID.

```
huawei(config)#ospfv3
huawei(config-ospfv3-1)#router-id 4.4.4.4
huawei(config-ospfv3-1)#quit
```

3. Enable OSPFv3 on the vlanif interface.

```
huawei(config)#interface vlanif 2
huawei(config-if-vlanif2)#ospfv3 1 area 0
huawei(config-if-vlanif2)#quit
```

4. Save the data.

```
huawei(config)#save
```

----End

Result

Run the **display ipv6 routing-table** command and you can find the learnt route table. Hosts can communicate with each other.

Configuration File

Configuration on each MA5600T/MA5603T/MA5608T is similar. Take Access Node_A for example.

```
ipv6
vlan 2 smart
port vlan 2 0/19 0
interface vlanif 2
ipv6 enable
ipv6 address 192:1::1:1 64
quit
ospfv3
router-id 1.1.1.1
quit
interface vlanif 2
ospfv3 1 area 0
ospfv3 dr-priority 100
quit
save
```

16.10.7 References

The following table lists the references of this document.

Document No.	Document Name	Protocol Compliance
RFC 2740	OSPF for IPv6	Fully compliant.
RFC 4552	Authentication/Confidentiality for OSPFv3	Fully compliant.
RFC 5187	OSPFv3 Graceful Restart	Fully compliant.
RFC 5643	Management Information Base for OSPFv3	Partially compli

Document No.	Document Name	Protocol Compliance
		ant. The SET operation is not supported.
RFC 7166	Supporting Authentication Trailer for OSPFv3	Fully compliant.

16.11 BGP

Border Gateway Protocol (BGP) is an of dynamic routing protocol used between autonomous system (ASs). As an external gateway protocol (EGP), BGP is different from OSPF and RIP, because BGP is mainly used for controlling the route propagation and selecting the optimal route instead of discovering and calculating routes.

16.11.1 Introduction to BGP

Definition

NOTE

- If BGP and BGP4+ implement a feature in the same way, details are not provided in this chapter.
- For the route aggregation function, BGP supports both automatic aggregation and manual aggregation, whereas BGP4+ supports only manual aggregation.
- BGP does not support MP-BGP. BGP4+ supports MP-BGP.

Border Gateway Protocol (BGP) is a dynamic routing protocol used between autonomous systems (AS).

BGP-1 (defined in RFC 1105), BGP-2 (defined in RFC 1163), and BGP-3 (defined in RFC 1267) are three earlier-released versions of BGP. BGP exchanges the reachable inter-AS routes, establishes inter-AS paths, avoids routing loops, and applies routing policies between ASs.

The current BGP version is BGP-4 defined by RFC 4271.

As an exterior routing protocol on the Internet, BGP is widely used among Internet Service Providers (ISP).

BGP has the following characteristics:

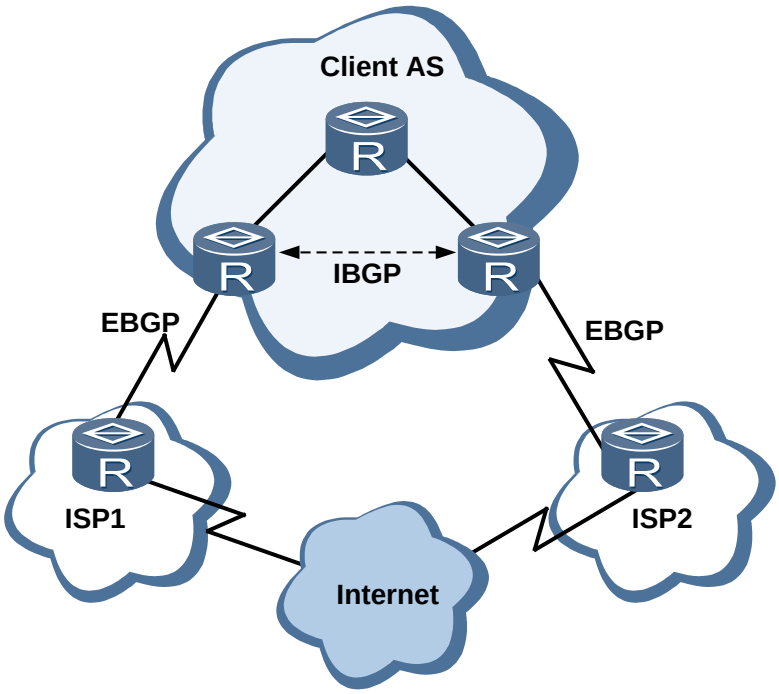
- Is an Exterior Gateway Protocol (EGP) unlike IGP such as OSPF and RIP, which controls route advertisement and selects the optimal route between ASs rather than discover and calculate routes.

- Uses the Transport Control Protocol (TCP) with the listening port number of 179 as the transport layer protocol.
  - BGP selects inter-AS routes, which imposes high requirements on the reliability of the protocol. TCP with high reliability is used to enhance the stability of BGP.
  - BGP peers must be logically connected and must establish TCP connections. The destination port number is 179 and the local port number is random.
- Supports Classless Inter-domain Routing (CIDR).
- Transmits only the updated routes when routes are being updated. This reduces the bandwidth occupied by BGP for route distribution. Therefore, BGP is applicable to the Internet where a large number of routes are transmitted.
- Is a distance-vector routing protocol.
- Is designed to avoid loops.
  - Inter-AS: BGP routes carry information about the ASs along the path. Routes that carry the local AS number are discarded, avoiding inter-AS loops.
  - Intra-AS: BGP does not advertise routes learned in the AS to the BGP peers, avoiding intra-AS loops.
- Provides rich routing policies to flexibly select and filter routes.
- Provides the mechanism for preventing route flapping, which effectively enhances the stability of the Internet.
- Can be easily extended to adapt to the development of networks.

Purpose

BGP transmits routes between ASs, but is not required in all situations.

Figure 16-65 BGP application scenario



BGP is required in the following situations:

- As shown in Figure 16-65, the user (Client AS) needs to be connected to two or more ISPs. The ISPs need to provide all or part of the Internet routes for the user. Based on the AS Path carried in BGP routes, the Router selects the optimal route through the AS of an ISP to the destination.
- Different organizations need to transmit the AS_Path.

BGP is not required in the following situations:

- User is connected to only one ISP.
- ISP does not need to provide Internet routes for users.
- ASs are connected through default routes.

16.11.2 Basic Principle of BGP

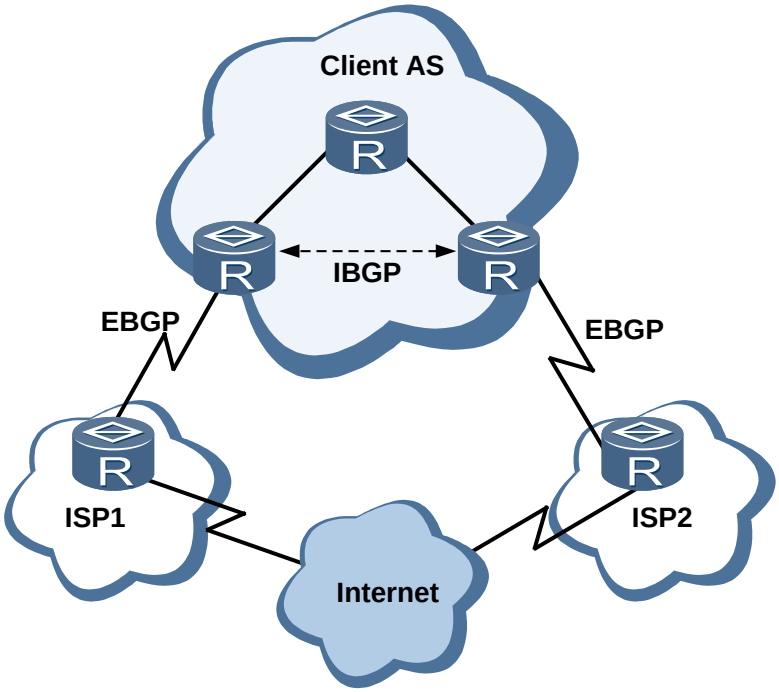
BGP Operating Modes

BGP operates on a Router in either of the following modes, as shown in Figure 16-66:

- Internal BGP (IBGP)
- External BGP (EBGP)

BGP is called IBGP when it runs within an AS It is called EBGP when it runs between ASs.

Figure 16-66 BGP operating modes



Roles in Transmitting BGP Messages

- Speaker: The device that sends BGP messages is called *a BGP speaker*. The speaker receives or generates new routing information, and then advertises the routing

information to other BGP speakers. When receiving a new route from another AS, a BGP speaker compares the route with the current route. If the route takes precedence over the existing route, or the route is new, the speaker advertises this route to all other BGP speakers except the BGP speaker that sent this route.

- Peer: BGP speakers that exchange messages with each other are called peers. Multiple peers compose a peer group.

BGP Messages

BGP runs by sending five types of BGP messages: Open, Update, Notification, Keepalive, and Route-refresh.

- Open message: is the first message that is sent after a TCP connection is set up, and is used to set up BGP peer relationships. After the peer receives an Open message and peer negotiation succeeds, the peer sends a Keepalive message to confirm and maintain the peer relationship. Then, peers can exchange Update, Notification, Keepalive, and Route-refresh messages.
- Update message: is used to exchange routes between BGP peers. Update messages can be used to send the following communications:
  - Advertise multiple reachable routes with the same attributes. These routes can share a group of route attributes. Route attributes contained in an Update message are applicable to all destination addresses (expressed by IP prefixes) contained in the Network Layer Reachability Information (NLRI) field of the Update message.
  - Withdraw multiple unreachable routes. Each route is identified by its destination address, which identifies routes previously advertised between BGP speakers.
  - Withdraw routes only. In this case, the message does not need to carry the path attributes or NLRI. Conversely, an Update message can be used only to advertise the reachable routes, so it does not need to carry information about withdrawn routes.
- Notification message: is sent to its peer when BGP detects an error. The BGP connection is then torn down immediately.
- Keepalive message: is sent periodically to the peer to maintain the peer relationship.
- Route-refresh message: is used to notify the peer of the capability to refresh routes.

  If all devices of BGP are enabled with Route-refresh capability, the local BGP device sends Route-refresh messages to peers when the import routing policy of BGP changes. After receiving the message, the peers resend their routing information to the local BGP device. The BGP routing table can be dynamically refreshed, and the new routing policy can be used, without tearing down BGP connections.

BGP Finite State Machine

The BGP Finite State Machine (FSM) has six states: Idle, Connect, Active, OpenSent, OpenConfirm, and Established.

- In Idle state, BGP denies all connection requests. This is the initial status of BGP.

  Upon receiving a Start event, BGP initiates a TCP connection to the remote BGP peer, starts the ConnectRetry Timer with the initial value, listens for a TCP connection initiated by the remote BGP peer, and changes its state to Connect.
- In Connect state, BGP performs other actions after TCP connection is set up
  - If the TCP connection succeeds, BGP stops the ConnectRetry Timer, sends an Open message to the remote peer, and changes its state to OpenSent.

  - If the TCP connection fails, BGP restarts the ConnectRetry Timer with the initial value, continues to listen for a TCP connection initiated by the remote peer, and changes its state to Active.
  - If the ConnectRetry Timer has expired before a TCP connection is established, BGP restarts the timer with the initial value, initiates a TCP connection to the remote BGP peer, and stays in the Connect state.
- In Active state, BGP attempts to set up a TCP connection. This is the intermediate status of BGP.
  - If the TCP connection succeeds, BGP stops the ConnectRetry Timer, sends an Open message to the remote peer, and changes its state to OpenSent.
  - If the ConnectRetry Timer has expired before a TCP connection is established, BGP restarts the timer with the initial value and changes its state to Connect.
  - If BGP initiates a TCP connection with an unknown IP address, the TCP connection fails. When this occurs, BGP restarts the ConnectRetry Timer with the initial value and stays in the Active state.
- In OpenSent state, BGP has sent one Open message to its peer and waits for the other Open message from the peer.
  - If there are no errors in the Open message received, BGP changes its state to OpenConfirm.
  - If there are errors in the Open message received, BGP sends a Notification message to the remote peer and changes its state to Idle.
  - If the TCP connection fails, BGP restarts the ConnectRetry Timer with the initial value, continues to listen for a TCP connection initiated by the remote peer, and changes its state to Active.
- In OpenConfirm state, BGP waits for a Notification message or a Keepalive message.
  - If BGP receives a Notification message or the TCP connection fails, BGP changes its state to Idle.
  - If BGP receives a Keepalive message, BGP changes its state to Established.
- In Established state, BGP peers can exchange Update, Route-Refresh, Keepalive, and Notification messages.

During establishment of BGP peer relationships, BGP is usually in the Idle, Active, or Established state.

- If BGP receives an Update or a Keepalive message, its state stays in Established.
- If BGP receives a Notification message, BGP changes its state to Idle.

The BGP peer relationship can be established only when both BGP peers are in the Established state. The two peers send Update messages to exchange routes.

BGP Processing

- BGP adopts TCP as its transport layer protocol. Before the BGP peer relationship is set up, a TCP connection must be set up between the peers. Then, BGP peers exchange Open messages to negotiate related parameters, and finally establish the BGP peer relationship.
- After the peer relationship is set up, BGP peers exchange BGP routing tables. BGP does not periodically update the routing table. When BGP routes change, however, BGP updates the BGP routing table incrementally through Update messages.
- BGP sends Keepalive messages to maintain the BGP connection between peers. When it detects an error on a network, for example, error packets or packets that indicate

unsupported negotiation capability are received, BGP sends a Notification message to report the error, and the BGP connection is torn down.

BGP Attributes

The BGP route attribute is a set of parameters that further describe routes. With the BGP route attribute, BGP can filter and select routes. BGP route attributes are classified into the following types:

- Well-known mandatory: can be identified by all BGP devices. This type of attribute is mandatory and must be carried in Update messages. Without this attribute, errors occur in the routing information.
- Well-known discretionary: can be identified by all BGP devices. The attribute is discretionary and is not necessarily carried in Update messages.
- Optional transitive: indicates the transitive attribute between ASs. A BGP device may not recognize this attribute, but it still receives these attributes and advertises them to other peers.
- Optional non-transitive: indicates an attribute that is not recognized. The corresponding attributes are ignored and are not advertised to other peers.

The common BGP route attributes are described as follows:

- Origin: defines the origin of a route and marks the paths of a BGP route. The Origin attributes are classified into the following types:
  - IGP: indicates the highest priority. For routing information obtained through an IGP of the AS that originates the route, the Origin attribute is IGP. For example, for routes imported to the BGP routing table through the **network** command, the Origin attribute is IGP.
  - Exterior Gateway Protocol (EGP): indicates the second highest priority. The Origin attribute of routes obtained through EGP is EGP.
  - Incomplete: indicates the lowest priority. The Origin attribute of routes learned by other means is Incomplete. For example, for the routes imported by BGP through the **import-route** command, the Origin attribute is Incomplete.
- AS_Path: is used to record all ASs that a route passes through from the local end to the destination in the distance-vector (DV) order.

  Assume that the BGP speaker advertises a local route:
  - When advertising the route to other ASs, the BGP speaker adds the local AS number in the AS_Path list, and advertises it to the neighboring devices through Update messages.
  - When advertising the route to the local AS, the BGP speaker creates an empty AS_Path list in an Update message.

  Assume that the BGP speaker advertises the routes learned from Update messages of other BGP speakers:
  - When advertising the route to other ASs, the BGP speaker adds the local AS number to the beginning of the AS_Path list. According to the AS_Path attribute, the BGP device that receives the route can detect the ASs through which the route passes to the destination. The number of the AS that is nearest to the local AS is placed at the top of the list. The other AS numbers are arranged in sequence.
  - When the BGP speaker advertises the route to the local AS, it does not change the AS_Path.

  The AS_Path attribute has four types:

  - AS_Sequence: a sequenced set of numbers of the ASs that a route passes through from a local end to the destination
  - AS_Set: an unsequenced set of numbers of the ASs that a route passes through from a local end to the destination The AS_Set attribute is used in the route aggregation scenario. After route aggregation, the device cannot sequence the numbers of ASs that specific routes pass through, so the AS_Set attribute is used to record the unsequenced AS numbers. No matter how many AS numbers an AS_Set contains, BGP regards the AS_Set as one AS number to calculate routes.
  - AS_Confed_Sequence: a sequenced set of sub-AS numbers in a confederation
  - AS_Confed_Set: an unsequenced set of sub-AS numbers in a confederation. The AS_Confed_Set attribute is used in the route aggregation scenario in a confederation.

  The AS_Confed_Sequence and AS_Confed_Set attributes are used to prevent route loops and to select routes among the various sub-ASs in a confederation.
- Next_Hop: is different from that of IGP. It is not necessarily the IP address of a neighboring device. Generally, the Next_Hop attribute complies with the following principles:
  - When advertising a route to an EBGP peer, the BGP speaker sets the next hop of the route to the address of the local interface through which the BGP peer relationship is set up.
  - When advertising a locally generated route to an IBGP peer, the BGP speaker sets the next hop of the route to the address of the local interface through which the BGP peer relationship is set up.
  - When advertising a route learned from an EBGP peer to an IBGP peer, the BGP speaker does not change the next hop of the route.
- Multi_Exit Discriminator (MED): is exchanged only between two neighboring ASs. The AS that receives the MED does not advertise it to any other ASs.

  MED serves as the metric used by an IGP. It is used to determine the optimal route when traffic enters an AS. When a BGP device obtains multiple routes to the same destination address but with different next hops through EBGP peers, the route with the smallest MED value is selected as the optimal route.
- Local_Pref: indicates preferences of the BGP devices. It is exchanged only between IBGP peers and is not advertised to other ASs.

  The Local_Pref attribute is used to determine the optimal route when traffic leaves an AS. When a BGP device obtains multiple routes to the same destination address but with different next hops through IBGP peers, the route with the largest Local_Pref value is selected.

Policies for BGP Route Selection

When there are multiple routes to the same destination, BGP selects routes according to the following policies:

1. Prefers the route with the highest PreVal.

   PrefVal is a Huawei-specific parameter. It is valid only on the device where it is configured.
2. Prefers the route with the highest Local_Pref.

   A route without Local_Pref has had the value set using the **default local-preference** command or has a value of 100 by default.

3. Prefers a locally originated route. A locally originated route takes precedence over a route learned from a peer.

   Locally originated routes include routes imported using the **network** command or the **import-route** command, manually summarized routes, and automatically summarized routes.

   a. A summarized route is preferred. A summarized route takes precedence over a non-summarized route.

   b. A route obtained using the **aggregate** command is preferred over a route obtained using the **summary automatic** command.

   c. A route imported using the **network** command is preferred over a route imported using the **import-route** command.

4. Prefers a route that carries the Accumulated Interior Gateway Protocol Metric (AIGP) attribute.

   - The priority of a route that carries the AIGP attribute is higher than the priority of a route that does not carry the AIGP attribute.
   - If two routes both carry the AIGP attribute, the route with a smaller AIGP attribute value plus IGP metric of the iterated next hop is preferred over the other route.

5. Prefers the route with the shortest AS_Path.

   - The AS_CONFED_SEQUENCE and AS_CONFED_SET are not included in the AS_Path length.
   - An AS_SET counts as 1, no matter how many ASs are in the set.
   - After you run the **bestroute as-path-ignore** command, the AS_Path attributes of routes are not compared in the route selection process.

6. Prefers the route with the highest Origin type. IGP is higher than EGP, and EGP is higher than Incomplete.

7. Prefers the route with the lowest MED.

   - BGP compares MEDs of only routes from the same AS, but not a confederation sub-AS. MEDs of two routes are compared only when the first AS number in the AS_SEQUENCE (excluding AS_CONFED_SEQUENCE) is the same for the two routes.
   - A route without MED is assigned a MED of 0, unless the **bestroute med-none-as-maximum** command is run. If you run the **bestroute med-none-as-maximum** command, the route is assigned the highest MED of 4294967295.
   - After you run the **compare-different-as-med** command, MEDs in routes received from peers in different ASs are compared. Do not use this command unless you confirm different ASs use the same IGP and route selection mode. Otherwise, a loop can occur.
   - If you run the **bestroute med-confederation** command, MEDs are compared for routes that consist of only AS_CONFED_SEQUENCE. The first AS number in the AS_CONFED_SEQUENCE must be the same for the routes.
   - After you run the **deterministic-med** command, routes are not selected in the sequence in which routes are received.

8. Prefers EBGP routes over IBGP routes.

   EBGP is higher than IBGP, IBGP is higher than LocalCross, and LocalCross is higher than RemoteCross.

   If the export route target (ERT) of a VPNv4 route in the routing table of a VPN instance on a Provide Edge (PE) matches the import route target (IRT) of another VPN instance

on the PE, the Virtual Private Network version 4 (VPNv4) route is added to the routing table of the second VPN instance. This is called LocalCross. If the ERT of a VPNv4 route from a remote PE is learned by the local PE and matches the IRT of a VPN instance on the local PE, the VPNv4 route will be added to the routing table of that VPN instance. This is called RemoteCross.

9. Prefers the route with the lowest IGP metric to the BGP next hop.

   After the **bestroute igp-metric-ignore** command is run, the IGP metrics are not compared for routes during route selection.

NOTE

Assume that load balancing is configured. If the preceding rules are the same and there are multiple external routes with the same AS_Path, load balancing will be performed based on the number of configured routes.

10. Prefers the route with the shortest Cluster_List.

NOTE

By default, Cluster_List takes precedence over Originator_ID during BGP route selection. To enable Originator_ID to take precedence over Cluster_List during BGP route selection, run the **bestroute routerid-prior-clusterlist** command.

11. Prefers the route advertised by the device with the smallest router ID.

NOTE

If routes carry the Originator_ID, the originator ID is substituted for the router ID during route selection. The route with the smallest Originator_ID is preferred.

12. Prefers the route learned from the peer with the smallest address if the IP addresses of peers are compared in the route selection process.

BGP ECMP

When multiple equal-cost routes have the same destination address, traffic can be evenly load balanced using BGP Equal Cost Multiple Path (ECMP).

Condition for BGP ECMP: Routes must have the same first nine attributes defined in the preceding "Policies for BGP Route Selection".

Policies for BGP Route Advertisement

BGP adopts the following policies for the BGP speaker to advertise routes:

- Advertises only the optimal route to its peer when there are multiple valid routes.
- Advertises the routes learned from EBGP devices to all BGP peers, including EBGP peers and IBGP peers.
- Does not advertise the routes learned from IBGP devices to its IBGP peers.
- Advertises the routes learned from IBGP devices to its EBGP peers.
- Advertises all BGP optimal routes to new peers when the peer relationship is established.

Synchronization of IBGP and IGP

IBGP and IGP are synchronized to prevent unreachable routes being imported to the external AS devices.

If a non-BGP device in an AS provides forwarding service, IP packets forwarded by this AS might be discarded because the destination address is unreachable. As shown in Figure 16-67, Router E learns route 8.0.0.0/8 of Router A from Router D through BGP, and then forwards the packet to Router D. Router D searches the routing table and detects that the next hop is

Router B. Router D forwards the packet to Router C through route iteration, because Router D obtained a route to Router B through IGP. Router C, however, does not obtain the route to 8.0.0.0/8 and discards the packet.

Figure 16-67 IBGP and IGP synchronization

If synchronization is configured, devices check the IGP routing table before they add the IBGP route to the routing table and advertising it to the EBGP peers. The IBGP route is added to the routing table and advertised to EBGP peers only when IGP obtains this IBGP route.

The synchronization can be disabled in the following cases:

- The local AS is not a transitive AS. (AS20 in Figure 1 is a transitive AS)
- All devices in the local AS are full-meshed IBGP peers.

16.11.3 Route Import

BGP itself cannot discover routes. It needs to import other protocol routes, such as IGP or static routes to the BGP routing table. In this manner, imported routes can be transmitted within an AS or between ASs.

BGP can import routes in either Import mode or Network mode.

- In Import mode, BGP imports routes according to protocol types, for example, Routing Information Protocol (RIP) routes, Open Shortest Path First (OSPF) routes, Intermediate System-to-Intermediate System (IS-IS) routes, static routes, or direct routes.
- The Network mode is more precise than the Import mode. In Network mode, routes with the specified prefix and mask are imported to the BGP routing table.

16.11.4 Route Summarization

Purpose

On medium or large-scale Border Gateway Protocol (BGP) networks, the BGP routing table on a device contains a large number of routing entries. Storing the routing table consumes a great deal of memory, and transmitting and processing routing information consume significant network resources. Route summarization can reduce the size of a routing table, prevent specific routes from being advertised, and minimize the impact of route flapping on network performance.

Definition

Route summarization is the process of summarizing specific routes with the same IP prefix into a summary route. BGP supports automatic and manual route summarization. Table 16-37 defines the differences between the two modes.

Table 16-37 Differences between automatic and manual route summarization

Route Summarization Mode	Implementation	Characteristics
Automatic route summarization	After automatic route summarization is configured, BGP summarizes routes based on the natural network segment and sends only the summarized route to peers. For example, 10.1.1.1/24 and 10.2.1.1/24 are summarized into 10.0.0.0/8, which is a Class A address.	• BGP summarizes only local routes that are imported using the **import-route** command. • During BGP route selection, an automatically summarized route has a lower priority than a manually summarized one. • An automatically summarized route does not carry path information because BGP summarizes only local routes that are imported using the **import-route** command. • BGP4+ does not support automatic route summarization.
Manual route summarization	BGP routes are summarized manually.	• The attributes carried in a manually summarized route are controllable. • Whether to advertise the specific routes for summarization is controllable. • During BGP route selection, a manually summarized route has a higher priority than an automatically summarized one. • A manually summarized route can carry specific path information, which prevents routing loops. • Both BGP and BGP4+ support manual route summarization.

An automatically summarized route comes from local routes, and the mechanism of automatic route summarization is much less complex than that of manual route summarization. Therefore, the next section describes only manual route summarization.

Related Concepts

Atomic_Aggregate: a well-known discretionary BGP attribute, carried in Update messages, indicating that a route is a summarized one. BGP speakers cannot delete this attribute during route transmission.

Aggregator: an optional transitive attribute, carried in Update messages, indicating where routes are summarized. **Aggregator** consists of the AS number and router ID of the router that performs the route summarization.

AS_Sequence: a type of **AS_Path**, carried in Update messages, recording in reverse order all the numbers of the ASs that a route passes from the local device to the destination address.

AS_Set: a type of **AS_Path**, carried in Update messages, recording all the numbers of the ASs that a route passes from the local device to the destination address without an order. **AS_Set** can also indicate a summarized route and carry path information. Therefore, if a summarized route carries **AS_Path**, **Atomic_Aggregate** is optional. During route selection, a router considers that **AS_Set** carries only one AS number regardless of the actual number of ASs.

NOTE

AS_Set affects BGP route selection. Whenever **AS_Set** changes, a router sends Update messages to its peers whose routes are not summarized by the router to notify the change. If the summarized route passes through a large number of ASs and the specific routes change frequently, the router needs to send Update messages frequently to its peers to notify them of the **AS_Set** changes. This process may lead to route flapping.

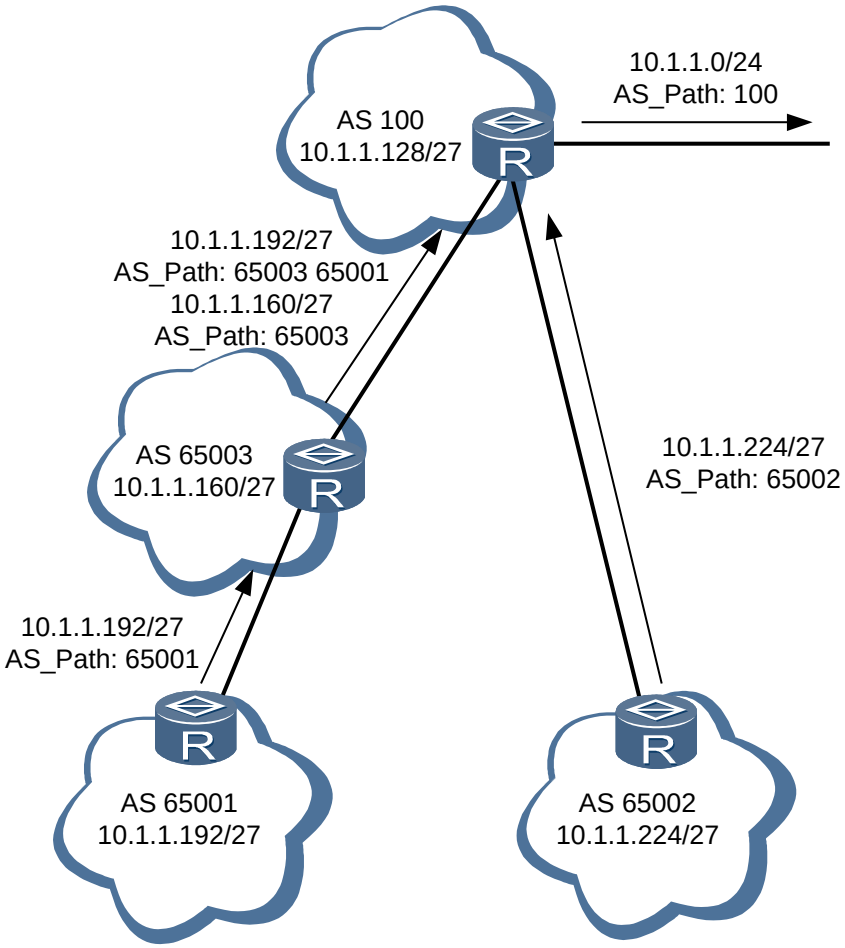
AS4_Path: a new attribute defined by BGP. It is similar to **AS_Path** in function, but **AS4_Path** can carry both 2-byte and 4-byte AS numbers. **AS4_Path** can be classified as **AS4_Sequence** or **AS4_Set**, which are respectively similar to **AS_Sequence** and **AS_Set** in function.

AS4_Aggregator: a new attribute defined by BGP. **AS4_Aggregator** carries 4-byte AS numbers, while **Aggregator** carries 2-byte AS numbers.

Implementation

As shown in Figure 16-68, the router in AS 100 summarizes the routes from AS 65001, AS 65002, and AS 65003 into the route 10.1.1.0/24 and then advertises it. Because the route 10.1.1.0/24 originates from AS 100, it carries only AS 100 without the path information about the specific routes for the summarization.

Figure 16-68 Networking for route summarization



Without the path information, **AS_Path** carried in the route 10.1.1.0/24 can no longer prevent routing loops. To warn downstream routers that the path information has been lost, the router in AS 100 adds Atomic_Aggregate to an Update message.

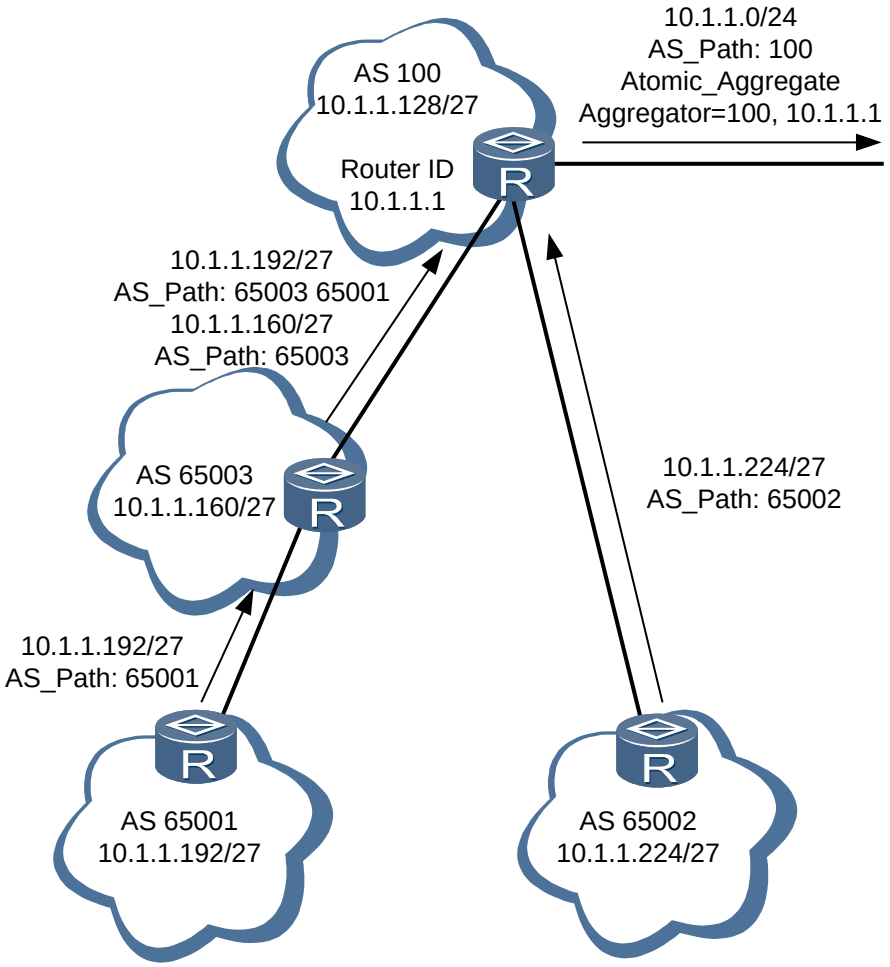
As shown in Figure 16-69, the router in AS 100 adds Atomic_Aggregate and Aggregator to an Update message to advertise the route 10.1.1.0/24.

NOTE

If Atomic_Aggregate is added to the route, Aggregator is optional.

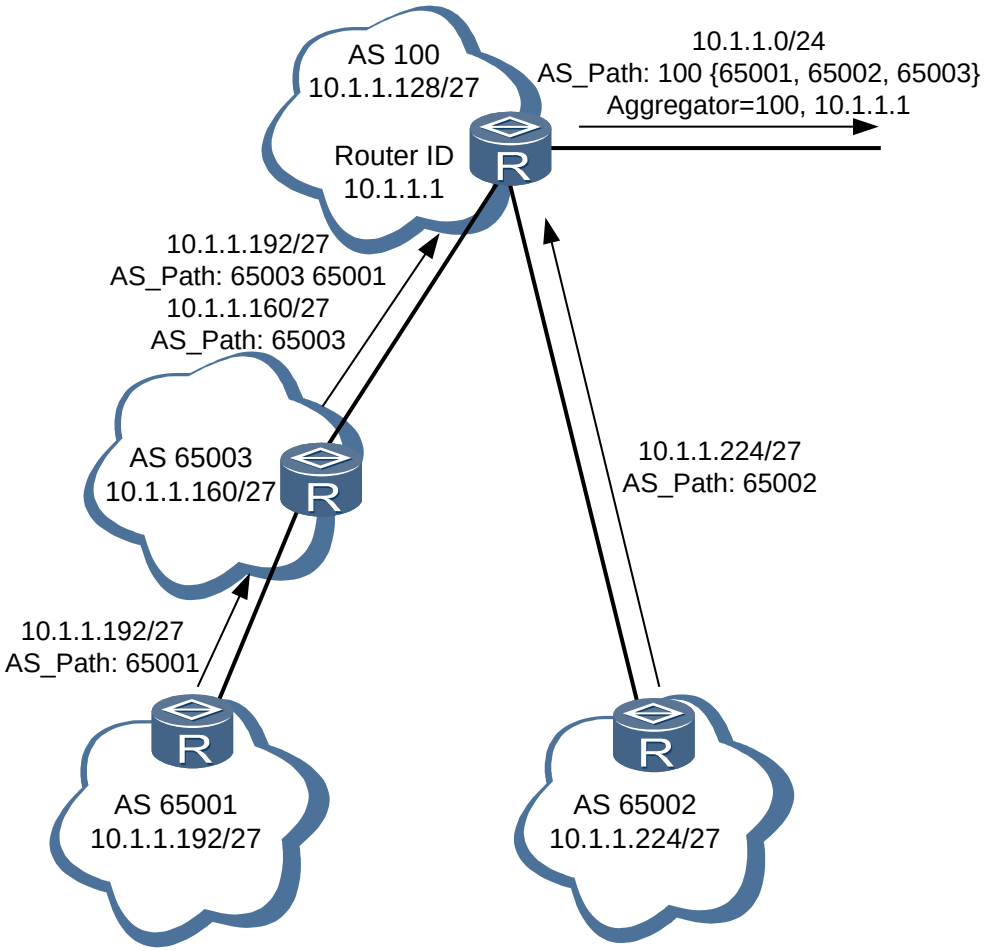
BGP speakers cannot delete **Atomic_Aggregate** carried in a summarized route during route transmission. After the downstream router receives this route, the router cannot restore the lost path information.

Figure 16-69 Networking in which an Update message carries **Atomic_Aggregate**



However, only Atomic_Aggregate and Aggregator cannot prevent routing loops. AS_Set can address this problem. If AS_Set is configured on the Router in AS 100 in the networking shown in Figure 16-70, the summarized route 10.1.1.0/24 carries AS_Set {65001, 65002, 65003} which records all the ASs it passes through.

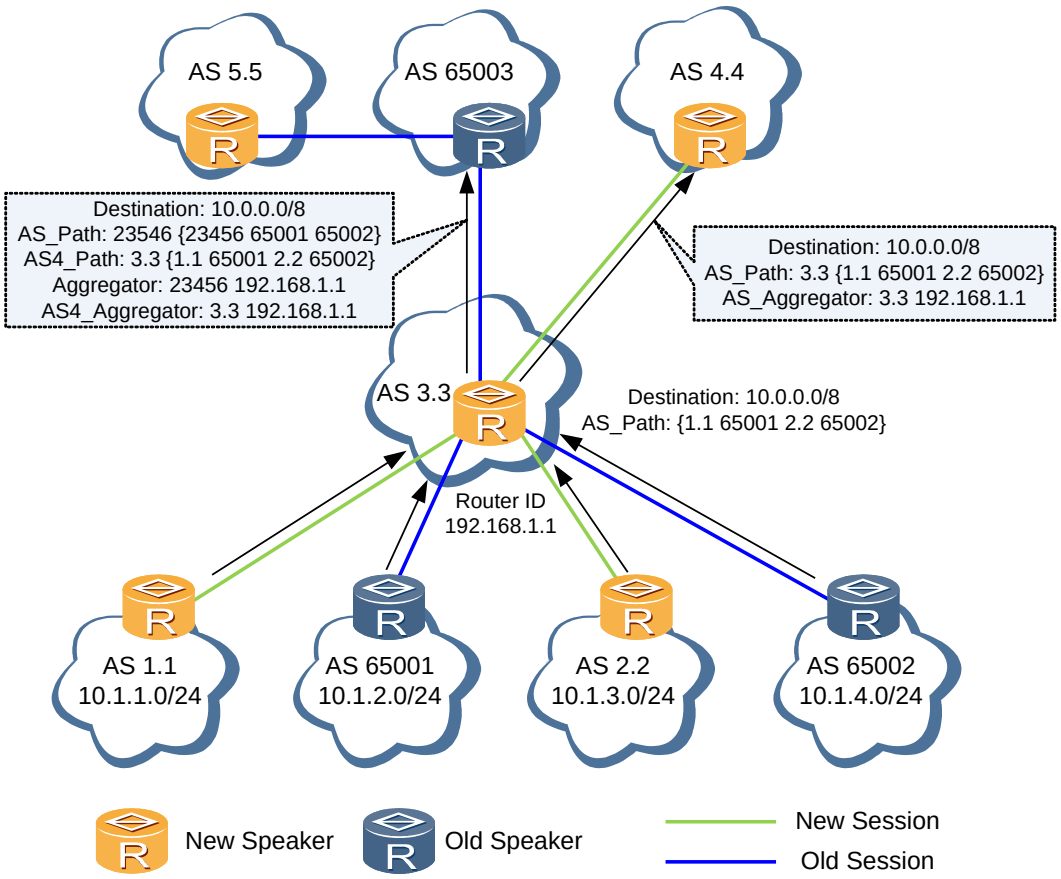
Figure 16-70 Networking in which an Update message carries AS_Set



Because the Router in AS 100 cannot determine the AS sequence, it records the ASs without an order in AS_Set. After the Router in AS 65001, AS 65002, or AS 65003 receives the route 10.1.1.0/24 elsewhere, it checks the **AS_Path** carried in the route. Because the **AS_Path** contains its AS number, the Router discards the route. Therefore, even though the ASs are listed without an order in AS_Set, routing loops can still be prevented.

In the preceding section, all Routers are old speakers. If new speakers co-exist with old speakers, AS4_Path and AS4_Aggregator must be available. Figure 16-71 shows such a scenario in which new speakers use 4-byte AS numbers and old speakers use 2-byte AS numbers. The Router in AS 3.3 summarizes the routes from AS 1.1, AS 65001, AS 2.2, and AS 65002 into the route 10.0.0.0/8 carrying **AS4_Path** {1.1, 65001, 2.2, 65002} and advertises it to the Routers in AS 4.4 and AS 65003. The carried **AS4_Path** equals an **AS4_Set** in function.

Figure 16-71 Networking for route summarization in which new speakers co-exist with old speakers



- Because the BGP connection is an old session between the Router in AS 3.3 and that in AS 65003 that does not support 4-byte AS numbers, the Router in AS 3.3 replaces the 4-byte AS numbers in **AS4_Path** and **AS4_Aggregator** with 23456 (**AS_Trans**) before it sends the route 10.0.0.0/8 to the Router in AS 65003. Therefore, the **AS_Path** carried in the route is 23456{23456, 65001, 23456, 65002}, and the **Aggregator** is 23456 192.168.1.1.

  After the Router in AS 65003 receives the route 10.0.0.0/8, it checks the **AS_Path**. Because its own AS number is not listed in the **AS_Path**, the Router in AS 65003 accepts the route.

NOTE

23456 is a reserved AS number and cannot be the number of the AS to which the downstream router that receives the summarized route belongs. Therefore, the downstream router does not discard the summarized route.

In addition, the Router in AS 65003 may be connected to downstream new speakers in an AS numbered in 4-byte format, AS 5.5 for example. To ensure that the Router in AS 5.5 knows about the actual path that the route passes through, the Router in AS 3.3 adds **AS4_Path** and **AS4_Aggregator** to the Update message to advertise the route 10.0.0.0/8 to the Router in AS 65003. After the Router in AS 65003 receives the message, it transparently transmits the message to the Router in AS 5.5. After the Router in AS 5.5 receives the message, it constructs the actual path that the route passes based on **AS4_Path** and **AS4_Aggregator** carried in the message.

- Because the BGP connection is a new session between the Router in AS 3.3 and that in AS 4.4 that supports 4-byte AS numbers, the Router in AS 3.3 adds only **AS_Path** and **AS_Aggregator** to the Update message to advertise the route 10.0.0.0/8 to the Router in AS 4.4. After the Router in AS 4.4 receives the route 10.0.0.0/8, it checks the **AS_Path**. Because its own AS number is not listed in the **AS_Path**, the Router in AS 4.4 accepts the route.

Benefits

Route summarization brings the following benefits:

- Reduces the router load: Route summarization reduces the size of a routing table and spares a router from advertising a large number of specific routes, which reduces the transmitting load. Route summarization also reduces the receiving load because downstream routers receive only the summarized route.
- Reduces the link load: A router advertises only the summarized route to its peers, which reduces link bandwidth consumption.
- Minimizes the impact of route flapping: If route flapping occurs in the ASs that the specific routes for summarization pass through, its impact will not spread beyond the ASs.

16.11.5 Route Dampening

Route instability is reflected in route flapping when a route in a routing table disappears and reappears frequently.

NOTE

A route is added to the routing table, and then is withdrawn. This occurrence is called route flapping.

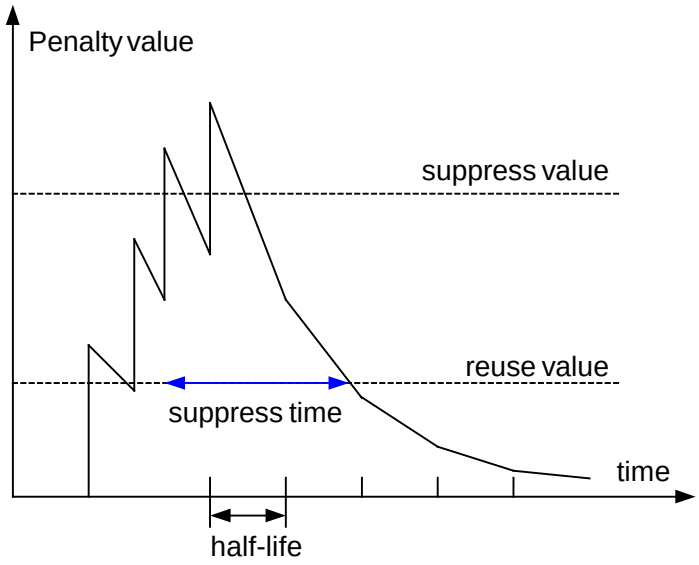
When route flapping occurs, a device sends a routing update to neighbors. The devices that receive the routing update need to recalculate routes and modify routing tables. Frequent route flapping consumes lots of bandwidth and CPU resources, and even affects the normal operation of a network.

Route dampening solves the problem of route instability caused by route flapping. In most situations, BGP is applied to complex networks where routes change frequently. To avoid the impact of frequent route flapping, BGP adopts route dampening to suppress unstable routes.

BGP dampening use a penalty value to measure the stability of a route. The greater the penalty value, the more unstable the route. Each time route flapping occurs (a route changes from active to inactive), BGP adds a certain penalty value (1000) to this route. When the penalty value of a route exceeds the suppression threshold, the route is suppressed. BGP does not add the route to the routing table, or advertise any Update message to BGP peers.

The penalty value of the suppressed route decreases to half after a certain period called *half life*. When the penalty value decreases to the reuse value, the route is reusable and is added to the routing table. At the same time, BGP advertises an Update message to BGP peers. The penalty value, suppression threshold, and half-life can be manually configured. Figure 16-72 shows the process of BGP route dampening.

Figure 16-72 Networking for BGP route dampening



16.11.6 Community Attribute

A community is a set of destination addresses with the same characteristics. The community attribute is expressed as a list in units of four bytes. The community is in the format of aa:nn or the community number.

- aa:nn: values range from 0 to 65535. The administrator can set a specific value as required. aa indicates the AS number and nn indicates the community identifier defined by the administrator. For example, if a route is from AS 100, and its community identifier is defined as 1, the format of the community is 100:1.
- Community number: is an integer that ranges from 0 to 4294967295. As defined in RFC 1997, the numbers from 0 (0x00000000) to 65535 (0x0000FFFF) and from 4294901760 (0xFFFF0000) to 4294967295 (0xFFFFFFFF) are reserved.

The community attribute is used to simplify the application, maintenance, and management of routing policies. With a community, a group of BGP devices in multiple ASs can share the same routing policy. Community is a route attribute that is transmitted between BGP peers and is not restricted by the AS. Before advertising a route with the community to peers, a BGP device can change the original community of the route.

The peer group allows a group of peers to share the same policy while the community allows a group of BGP routes to share the same policy.

Besides the well-known communities, you can define a community filter to filter extended communities to control routing policies in a more flexible manner.

Well-known Community

Table 16-38 lists the well-known community attributes of BGP routes.

Table 16-38 Well-known communities of BGP routes

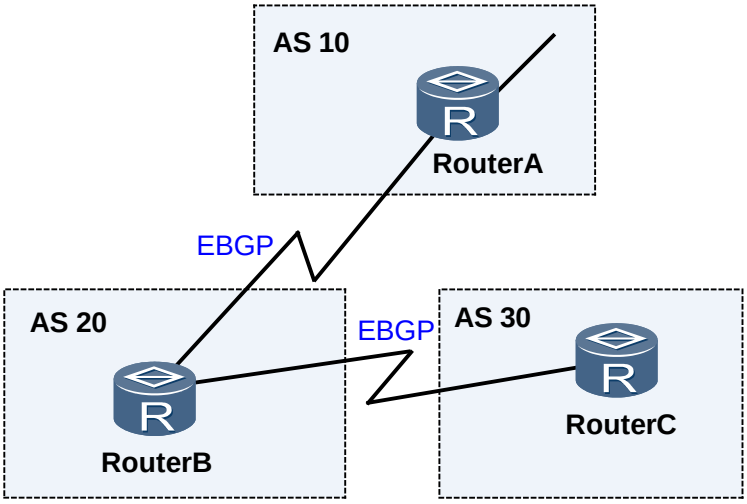
Community	Identifier	Description

Community	Identifier	Description
Internet	0 (0x00000000)	By default, all routes belong to the Internet community. A route with this attribute can be advertised to all BGP peers.
No_Export	4294967041 (0xFFFFFF01)	A route with this attribute cannot be advertised outside the local AS. If a confederation is defined, the route with this attribute cannot be advertised to ASs outside the confederation, but only to other sub-ASs in the confederation.
No_Advertise	4294967042 (0xFFFFFF02)	A route with this attribute cannot be advertised to any other BGP peers.
No_Export_Subconfed	4294967043 (0xFFFFFF03)	A route with this attribute cannot be advertised outside the local AS or to other sub-ASs in the confederation.

Networking Applications

In Figure 16-73, EBGP connections are established between Router B and Router A, and between Router B and Router C. With the community attribute of No_Export configured on Router A, routes from AS10 advertised to AS20 are not advertised to other ASs by AS20.

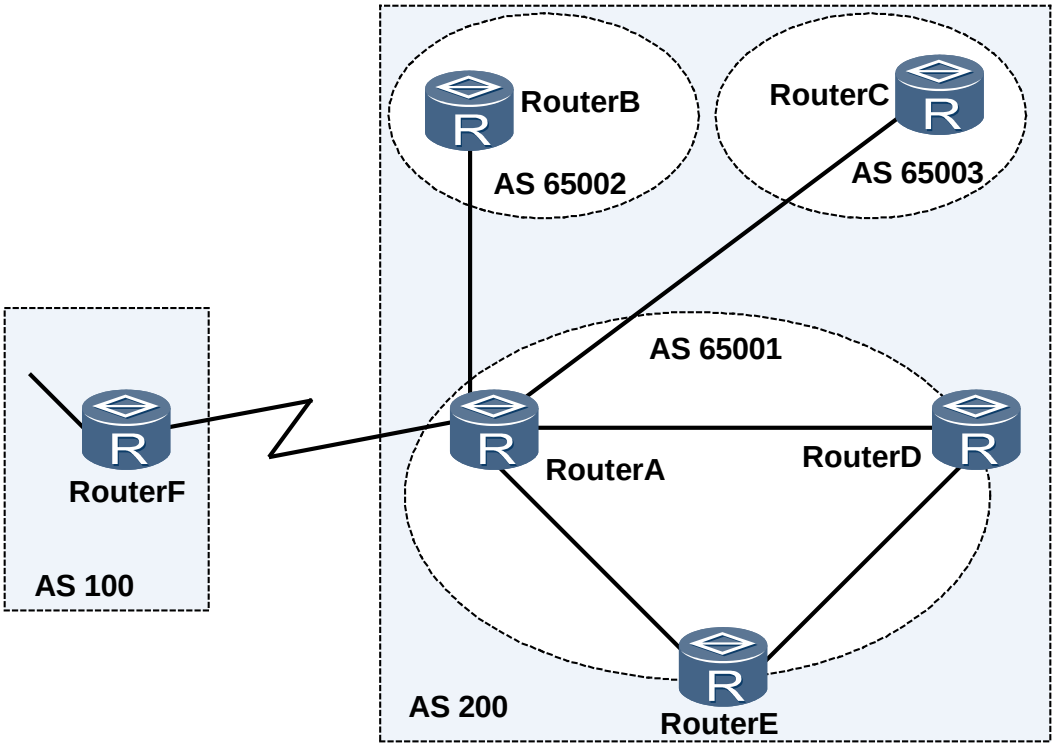
Figure 16-73 Networking for configuring BGP communities



16.11.7 BGP Confederation

The *confederation* is another method of dealing with increasing IBGP connections in an AS. It divides an AS into several sub-ASs. IBGP full meshes are established in each sub-AS, and EBGP full meshes are established between sub-ASs, as shown in Figure 16-74.

Figure 16-74 Confederation



For BGP speakers outside the confederation (such as, the devices in AS100), the sub-ASs (AS65001, AS65002, and AS65003) in the same confederation are invisible. The external devices do not need to detect the topology of each sub-AS. The confederation ID is the AS number that is used to identify the entire confederation. As shown in Figure 16-74, AS200 is the confederation ID.

As shown in Figure 16-74, AS200 has multiple BGP devices in AS200. To reduce IBGP connections, AS200 is divided into three sub-ASs: AS65001, AS65002, and AS65003. In AS65001, IBGP full meshes are established between the three devices.

Applications and Limitations

The confederation needs to be configured on each device, and any device that joins the confederation must support the confederation function.

The confederation has disadvantages. For example, if the devices are not in a confederation originally, but later need to be configured as a confederation, the logical topology changes accordingly.

In large-scale BGP networks, RR and confederation can be used together.

NOTE

The old BGP speaker with 2-byte AS numbers and the new speaker with 4-byte AS numbers cannot exist in the same confederation. Because AS4_Path does not support confederations, routing loops can occur.

16.11.8 MP-BGP and Address Families

The Border Gateway Protocol 4 (BGP-4) transmits only IPv4 unicast routing information and cannot transmit routing information for other network layer protocols, such as IPv6 and multicast protocols.

To support multiple types of network layer protocols, the Internet Engineering Task Force (IETF) extended BGP-4 to Multiprotocol Extensions for BGP-4 (MP-BGP). The current MP-BGP standard is RFC 4760. MP-BGP is forward compatible. Routers that support BGP extension can communicate with Routers that do not.

As an enhancement of BGP-4, MP-BGP provides routing information for various protocols, such as IPv6 (BGP4+) and multicast:

- MP-BGP maintains unicast and multicast routing information, and stores both types in different routing tables to ensure their separation.
- MP-BGP supports unicast and multicast, and constructs different network topologies for each.
- MP-BGP can maintain unicast and multicast routes based on routing policies. The unicast routing policies and configurations supported by BGP-4 can mostly be applied to multicast.

Extended Attributes

BGP-4 update packets carry three IPv4-related attributes: network layer reachable information (NLRI), Next_Hop, and Aggregator. Aggregator contains the IP address of the BGP speaker that performs route aggregation.

To support multiple types of network layer protocols, BGP-4 needs to carry the information about network layer protocols in NLRI and Next_Hop. MP-BGP introduces the following route attributes:

- MP_REACH_NLRI: indicates the multiprotocol reachable NLRI, which is used to advertise a reachable route and the next hop.
- MP_UNREACH_NLRI: indicates the multiprotocol unreachable NLRI, which is used to withdraw an unreachable route.

The two attributes are optional non-transitive. Therefore, BGP speakers that do not support multiprotocol ignore the information carried by the two attributes, and do not advertise the information to peers.

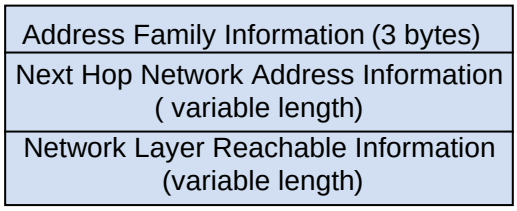
NOTE

Optional non-transitive is a BGP attribute type. If a BGP device does not support this attribute type, the Update messages with attributes of this type are ignored, and the messages are not advertised to other peers.

MP_REACH_NLRI

Multiprotocol Reachable NLRI (MP_REACH_NLRI) is used to advertise reachable routes and information about the next hop. MP_REACH_NLRI is coded as one or more 3-tuples of the form <Address Family Information, Next Hop Information, Network Layer Reachability Information>.

Figure 16-75 Format of the MP_REACH_NLRI field

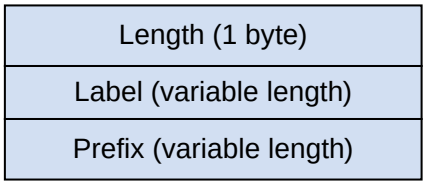


Descriptions of each part of the MP_REACH_NLRI field are as follows:

- Address Family Information: consists of a 2-byte Address Family Identifier (AFI) and a 1-byte Subsequent Address Family Identifier (SAFI):
  - The AFI identifies a network layer protocol. Defined values for this field are specified in RFC 1700 (Address Family Number). For example, 1 indicates IPv4; 2 indicates IPv6.
  - The SAFI indicates the type of the NLRI field.

  If the AFI is 1 and the SAFI is 128, the address in the NLRI field is a BGP-VPNv4 address.
- Next Hop Network Address Information: consists of the 1-byte length of the next-hop network address and the next-hop network address of a variable length. A next-hop network address refers to the network address of the next device on the path to the destination.
- Network Layer Reachable Information: is a variable-length field that lists NLRI for the routes being advertised. Figure 16-76 shows the format of the NLRI field.

  Figure 16-76 Format of the NLRI field that carries label information

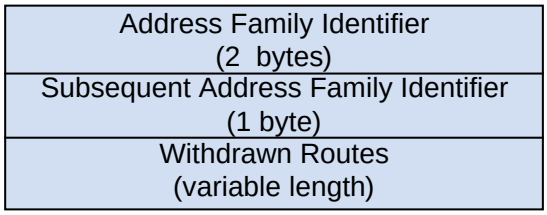
  Descriptions of each part of the NLRI field are as follows:
  - Length: indicates the total bits of the label and prefix.
  - Label: consists of one or more labels. The length of a label is 3 bytes. For more information about labels, see *MPLS*.
  - Prefix: In a BGP/Multiprotocol Label Switching (MPLS) IP VPN, the prefix field consists of a route distinguisher (RD) and IPv4 address prefix.

VPNv4 update messages exchanged between provider edges (PEs) or autonomous system boundary routers (ASBRs) carry MP_REACH_NLRI. An Update message can carry multiple reachable routes with the same routing attributes.

MP_UNREACH_NLRI

Multiprotocol Unreachable NLRI (MP_UNREACH_NLRI) is used to inform a peer to delete unreachable routes. Figure 16-77 shows the format of the attribute.

Figure 16-77 Format of the MP_UNREACH_NLRI field



Descriptions of each part of the MP_UNREACH_NLRI field are as follows:

- AFI: identifies a network layer protocol. AFI uses the address family values defined in RFC 1700.
- SAFI: indicates the NLRI type and is similar to SAFI in MP_REACH_NLRI.
- Withdrawn Routes: indicates an unreachable route list, which consists of one or more NLRI fields. In the Withdrawn Routes field, BGP speakers can withdraw the route by filling the NLRI field in the same manner as that for the previously advertised reachable route.

Update messages carrying MP_UNREACH_NLRI are sent to withdraw routes. An Update message can carry information about multiple unreachable routes.

If the labels of routes to be withdrawn are specified in the messages, the routes with specified labels are withdrawn. If the labels are not specified, only the routes without labels are withdrawn.

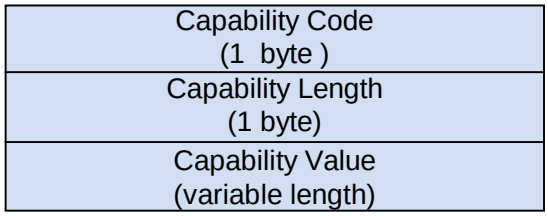
Update messages with MP_UNREACH_NLRI do not carry any path attributes.

Negotiation of the MP-BGP Capability

A BGP device gets to know the negotiation capability of its peer by checking the capability parameters in Open messages. If the BGP device and its peer support the same function, the BGP device and its peer communicate using the function.

The optional parameters of the negotiation capability in an Open message consist of three parts: Capability Code, Capability Length, and Capability Value. Figure 16-78 shows the format of the capability parameters.

Figure 16-78 Format of BGP capability parameters

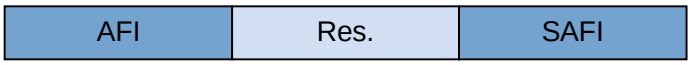


Descriptions of BGP capability parameters are as follows:

- Capability Code: uniquely identifies the capability type. The value 1 indicates that the BGP speaker has the MP-BGP capability.

- Capability Length: indicates the length of the capability field. For MP-BGP, the length of the capability field is 4.
- Capability Value: indicates the value of the capability field. The length is variable and depends on the type specified in Capability Code. Figure 16-79 shows the format of the Capability Value field:
  - The meanings of 2-byte AFI and 1-byte SAFI are the same as those of MP_REACH_NLRI.
  - Res. is a 1-byte reserved field. The sender sets the value to 0, and the receiver ignores the field.

Figure 16-79 Format of the Capability Value field



At present, BGP does not support dynamic capability negotiation. After a BGP speaker advertises an Open message with optional capability fields,

- If the speaker receives a Notification message from its peer, the peer does not support the capability. Then the BGP speaker tears down the session with its peer and sends an Open message without any optional capability fields to the peer to attempt a new BGP connection.
- If the peer supports the capability advertisement but the capability fields are unknown or unsupported, negotiation fails. Then the BGP speaker tears down the session with its peer, and sends an Open message without the optional capability fields (but may carry other optional capability fields) to the peer to attempt a new BGP connection.

After any change in BGP capabilities, such as enabling or disabling label-routing capabilities, enabling or disabling address family capabilities (IPv4, IPv6, VPNv4, and VPNv6), and enabling graceful restart (GR) capabilities, the BGP speaker tears down the session with its peer, and then re-negotiates the capabilities with its peer.

Address Family

BGP uses address families to differentiate network protocol applications. The MA5600T/MA5603T/MA5608T supports a wide range of MP-BGP applications. These applications can be configured in their respective extended BGP address family views.

Table 16-39 lists BGP address families.

Table 16-39 BGP address families

BGP Address Family	AFI	SAFI	Description
BGP-IPv4 unicast address family	1	1	Allows PEs to maintain public network BGP peer relationships and transmit public network IPv4 routing information.
BGP VPN instance IPv4	1	1	Allows BGP to transmit VPN IPv4 routing information between PEs and CEs in a BGP/MPLS IPv4 VPN scenario.

BGP Address Family	AFI	SAFI	Description
address family			
BGP-IPv6 unicast address family	2	1	Allows BGP to transmit IPv6 routing information on an IPv6 public network. The BGP-IPv6 unicast address family maintains IPv6 BGP peer relationships and transmits public network IPv6 routing information. The BGP-IPv6 unicast address family enables PEs to establish BGP 6PE peer relationships, allowing separate IPv6 networks to connect to an IPv4 public network.
BGP VPN instance IPv6 address family	2	1	Allows BGP to transmit VPN IPv6 routing information between PEs and CEs in a BGP/MPLS IPv6 VPN scenario.

16.11.9 BGP GR

Graceful restart (GR) is one of the high availability (HA) technologies, which comprise a series of comprehensive technologies such as fault-tolerant redundancy, link protection, faulty node recovery, and traffic engineering. As a fault-tolerant redundancy technology, GR ensures normal forwarding of data during the restart of routing protocols to prevent interruption of key services. Currently, GR has been widely applied to the master/slave switchover and system upgrade.

GR is usually used when the active route processor (RP) fails because of a software or hardware error, or used by an administrator to perform the master/slave switchover.

Prerequisite for Implementation

On a traditional routing device, a processor implements both control and forwarding. The processor finds routes based on routing protocols, and maintains the routing table and forwarding table of the device. Mid-range and high-end devices generally adopt the multi-RP structure to improve forwarding performance and reliability. The processor in charge of routing protocols is located on the main control board, whereas the processor responsible for data forwarding is located on the interface board. The design helps to ensure the continuity of packet forwarding on the interface board during the restart of the main processor. The technology that separates control from forwarding satisfies the prerequisite for GR implementation.

At present, a GR-capable device must have two main control boards. In addition, the interface board must have an independent processor and memory.

Related Concepts

The concepts related to GR are as follows:

- GR Restarter: indicates a device that performs master/slave switchover triggered by the administrator or a failure. A GR Restarter must support GR.

- GR helper: indicates the neighbor of a GR Restarter. A GR helper must support GR.
- GR session: indicates a session, through which a GR Restarter and a GR helper can negotiate GR capabilities.
- GR time: indicates the time when the GR helper finds that the GR Restarter is Down but keeps the topology information or routes obtained from the GR Restarter.
- End-of-RIB (EOR): indicates a BGP information, notifying a peer BGP that the first route upgrade is finished after the negotiation.
- EOR timer: indicates a maximum time of a local device waiting for the EOR information sent from the peer. If the local device does not receive the EOR information from the peer within the EOR timer, the local device will select an optimal route from the current routes.

Principles

Principles of BGP GR are as follows:

- During BGP peer relationship establishment, devices negotiate GR capabilities by sending supported GR capabilities to each other.
- When detecting the master/slave switchover of the GR Restarter, a GR helper does not delete the routing information and forwarding entries related to the GR Restarter within the GR time, but waits to re-establish a BGP connection with the GR Restarter.

NOTE

If the GR Helper sends Keepalive packets to the Restarter but receives no reply within the Holdtimer, the GR Helper is in GR state and marks the route sent from the Restarter as Stale. The Restarter restart may trigger the GR Helper to enter the GR state.

The **reset bgp** command is used to restart BGP without triggering the BGP GR.

- After the master/slave switchover, the GR Restarter receives routes from all the negotiated peers with GR capabilities before the switchover, and starts the EOR timer. The GR Restarter selects a route when either of the following conditions is met:
  a. The GR Restarter receives the EOR information of all peers and the EOR timer is deleted.
  b. The EOR timer times out but the GR Restarter receives no EOR information from all peers.
- The GR Restarter sends the optimal route to the GR Helper and the GR Helper starts the EOR timer. The GR Helper quits GR when either of the following conditions is met:
  a. The GR Helper receives the EOR information from the GR Restarter and the EOR timer is deleted.
  b. The EOR timer times out and the GR Helper receives no EOR information from the GR Restarter.

GR Reset

Currently, BGP does not support dynamic capability negotiation. Therefore, each time a new BGP capability (such as the IPv4, IPv6, VPNv4, and VPNv6 capabilities) is enabled on a BGP speaker, the BGP speaker tears down existing sessions with its peer and renegotiates BGP capabilities. This process will interrupt ongoing services.

To prevent the service interruptions, the MA5600T/MA5603T/MA5608T provides the GR reset function that enables the MA5600T/MA5603T/MA5608T to reset a BGP session in GR mode. With the GR reset function configured, when you enable a new BGP capability on the BGP speaker, the BGP speaker enters the GR state, resets the BGP session, and renegotiates BGP capabilities with the peer. In the whole process, the BGP speaker re-establishes the

existing sessions but does not delete the routing entries for the existing sessions, so that the existing services are not interrupted.

Benefits

Through BGP GR, the forwarding is not interrupted. In addition, the flapping of BGP occurs only on the neighbors of the GR Restarter, and does not occur in the entire routing domain. This is important for BGP that needs to process a large number of routes.

16.11.10 BGP Dynamic Update Peer-Groups

With rapid increases in the size of the routing table and the complexity of the network topology, BGP needs to support more neighbors. Especially in the case of a large number of neighbors and routes, high-performance packaging and forwarding are required when a device needs to send routes to a large number of BGP neighbors, most of which share the same outbound policies.

The dynamic update peer-groups feature treats all BGP neighbors with the same outbound policies as an update-group.

Without the dynamic update peer-groups feature, each route to be sent is grouped per neighbor. With the dynamic update peer-groups feature, routes are grouped uniformly and sent separately. That is, each route to be sent is grouped once and then sent to all neighbors in the update-group, which improves grouping efficiency exponentially. When a large number of neighbors and routes exist, the BGP dynamic update peer-groups feature greatly improves BGP route packaging and forwarding performance.

Application Environment

The application scenarios of the BGP dynamic update peer-groups feature are as follows:

- International gateway
- RR
- A device sends the routes received from EBGP neighbors to all IBGP neighbors

  The following figures represent each scenario in turn.

Figure 16-80 Networking for the international gateway

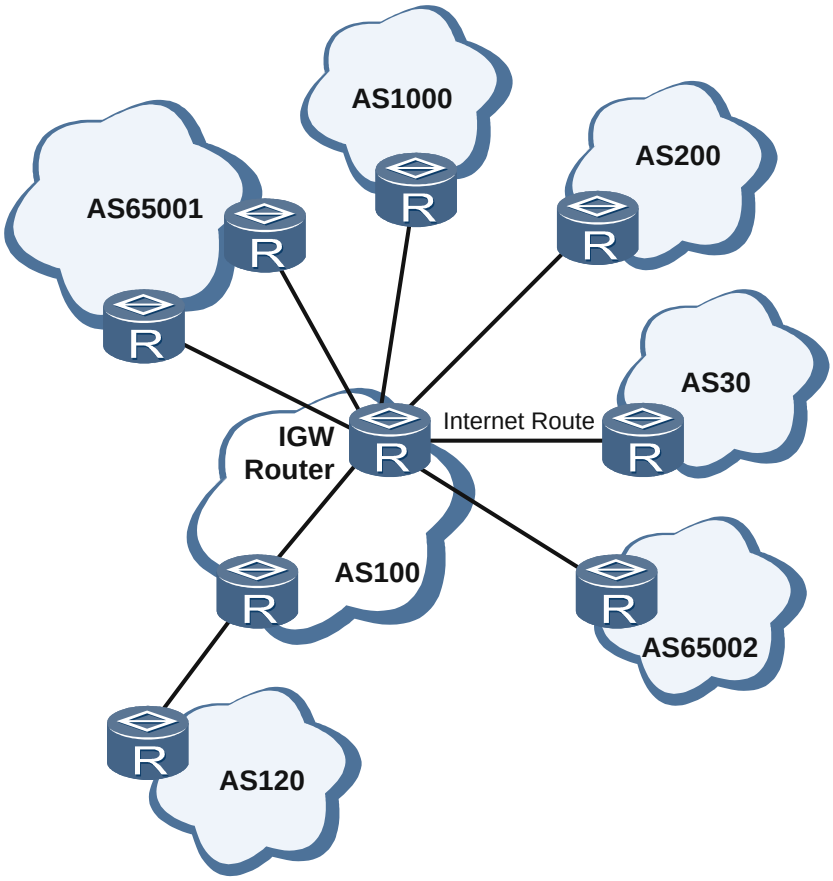


Figure 16-81 Networking for the RR with many clients

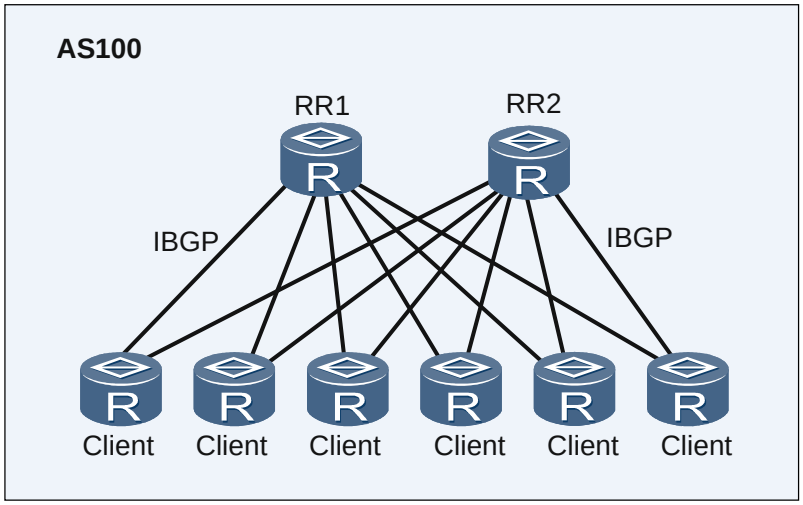
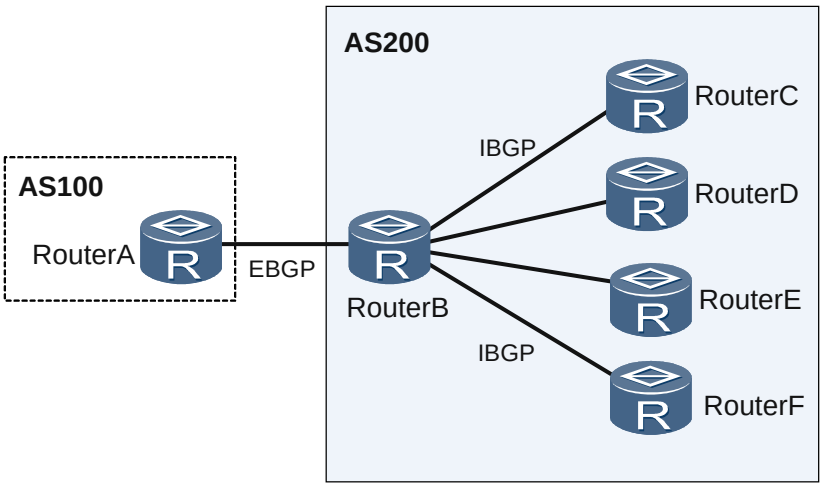


Figure 16-82 Networking for a PE connecting multiple IBGP neighbors



In the preceding scenarios, a device often needs to send routes to a large number of BGP neighbors, most of which share the same outbound policies. This situation is most obviously presented in the networking shown in Figure 16-81. When a large number of neighbors and routes exist, the forwarding efficiency is restricted.

After the dynamic update peer-groups feature is applied, each route to be sent is grouped once and sent to all neighbors in the update-group. For example, an RR has clients and needs to reflect routes to them. If the RR sends routes grouped per neighbor to 100 clients, the total number of times that all routes are grouped is . When the dynamic update peer-groups feature is applied, the total number of grouping times changes to . and performance is improved by a factor of .

16.11.11 4-Byte AS Number

Currently, 2-byte AS numbers used on the network range from 0 to 65535 and available AS numbers almost become exhausted. Therefore, 2-byte AS numbers need to be extended to 4-byte AS numbers, which are also compatible with the old speaker that supports only 2-byte AS numbers.

The 4-byte AS number feature extends a 2-byte AS number to a 4-byte AS number. The feature defines a new capability code and new optional transitive attributes t negotiate the 4-byte AS number capability and transmit 4-byte AS numbers. This implements communications between new speakers that support 4-byte AS numbers, and between old speakers that support only 2-byte AS numbers and new speakers.

- New speaker: indicates the peer that supports 4-byte AS numbers.
- Old speaker: indicates the peer that does not support 4-byte AS numbers.
- New session: indicates the BGP connection between new speakers.
- Old session: indicates the BGP connection between new speakers and old speakers, or between old speakers.

BGP Extension

An open capability code 0x41 is defined for BGP connection negotiation. The code 0x41 indicates that the BGP speaker supports 4-byte AS numbers.

Two new optional transitive attributes, AS4_Path with attribute code 0x11 and AS4_Aggregator with the attribute code 0x12, are defined to transmit 4-byte AS numbers on old sessions.

If a connection is set up between a new speaker and an old speaker and the AS number of the new speaker is greater than 65535, the remote AS number needs to be specified as AS_TRANS on the old speaker. AS_TRANS is a reserved AS number with the value 23456.

Principles

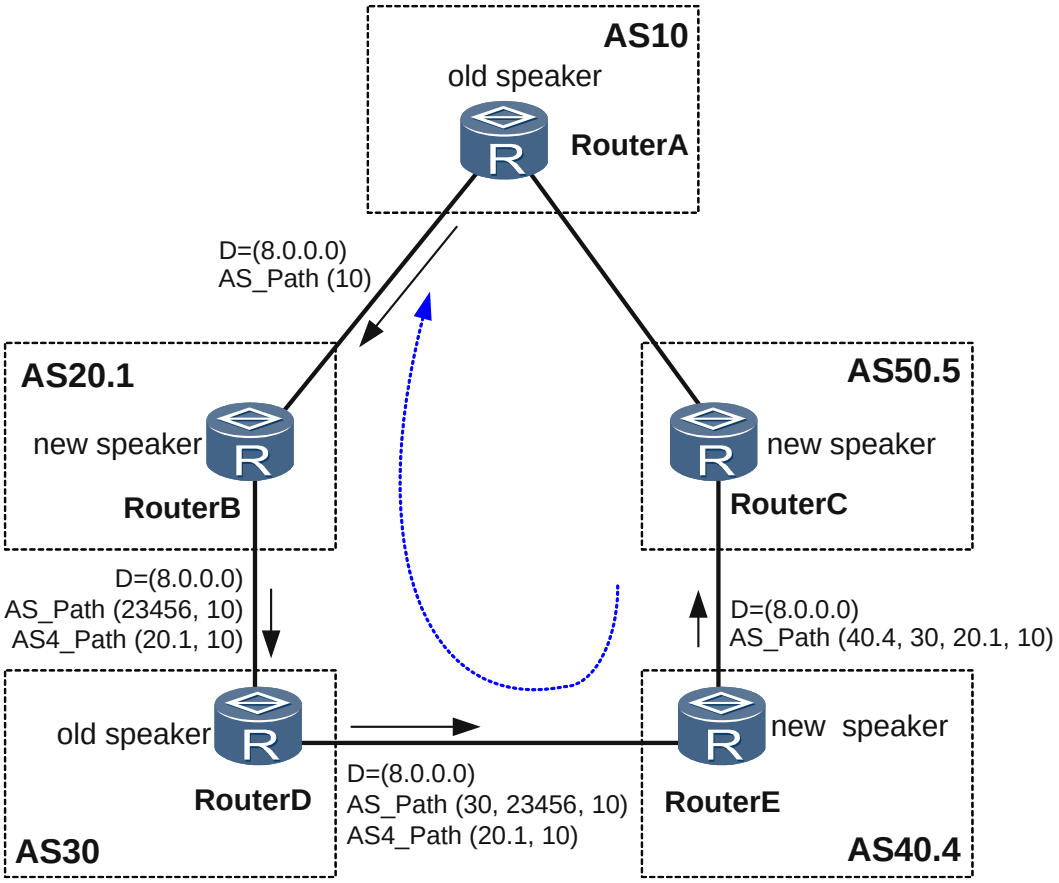
When setting up connections, BGP neighbors determine whether the peer supports 4-byte AS numbers according to the optional capability field in Open messages.

- New sessions are set up between new speakers. AS_Path and Aggregator in an Update message carry 4-byte AS numbers.
- Old sessions are set up between new and old speakers. AS_Path and Aggregator on old speakers carry 2-byte AS numbers.
  - When a new speaker sends an Update message to an old speaker, if the AS number of the new speaker is greater than 65535, AS4_Path and AS4_Aggregator are used together with AS_Path and AS_Aggregator to carry 4-byte AS numbers. AS4_Path and AS4_Aggregator are transparent to the old speaker.
  - When receiving messages that contain AS_Path, AS4_Path, AS_Aggregator, and AS4_Aggregator from an old speaker, a new speaker reconstructs the actual AS_Path and AS_Aggregator based on the reconstruction algorithm.

Application Environment

Figure 16-83 shows old speakers that support 2-byte AS numbers and new speakers that support 4-byte AS numbers in the topology. The 4-byte AS number feature, together with AS4_Path, transmits routing information between the old and new speakers.

Figure 16-83 Networking for the application of 4-byte AS numbers



As shown in Figure 16-83, before advertising route D=8.0.0.0 of AS 10 to other ASs, a BGP device performs the following:

1. BGP adds the AS number of AS10 to the AS_Path list (10).
2. When the route passes AS 20.1, to enable Router D (old speaker) to transmit AS path information with 4-byte AS numbers, this route carries the AS4_Path attribute (20.1, 10). Router B then adds the route AS number 20.1 to the beginning of the AS_Path list (23456, 10). (The value 23456 is obtained when AS_TRANS replaces 20.1.)
3. When the route passes AS30, Router D, an old speaker, transparently transmits AS4_Path (20.1, 10) to Router E. Router D then adds the route AS number 30 to the beginning of the AS_Path list (30, 23456, 10).
4. When the route passes AS 40.4, after the reconstruction of AS_Path and AS4_Path, BGP adds 40.4, the AS number of AS 40.4, to the beginning of the AS_Path list (40.4, 30, 20.1, 10).

   The rest may be deduced by analogy. After the device in AS 50.5 receives the route, the device learns the path to AS 10 according to the AS_Path list.

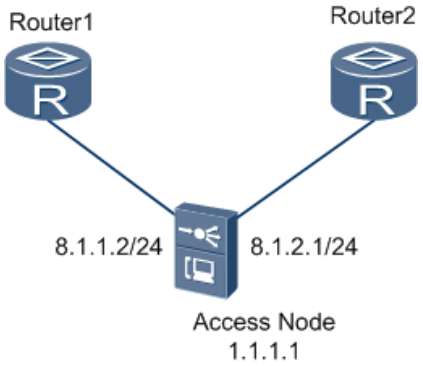
16.11.12 Configuration Example of BGP

This topic provides an example for configuring the BGP on the device.

Service Requirements

In this example network, an IBGP connection is set up between access node and routers.

Figure 16-84 Example network for configuring the BGP



Data Plan

Table 16-40 provides the data plan for configuring the BGP.

Table 16-40 Data plan for configuring the BGP

Item	Data	Remarks
Access node	IP address of VLAN interface 6: 8.1.1.2/24	It is used for the IBGP connection to the Router1.
	IP address of VLAN interface 2: 8.1.2.1/24	It is used for the IBGP connection to the Router12.
	Router ID: 1.1.1.1	-
	AS number: 2000	-

Procedure

Configure access node.

1. Configure the IP address of the Layer 3 interface.

```
huawei(config)#vlan 6 smart
huawei(config)#port vlan 6 0/19 0
huawei(config)#interface vlanif 6
huawei(config-if-vlanif6)#ip address 8.1.1.2 24
huawei(config-if-vlanif6)#quit
huawei(config)#vlan 2 smart
huawei(config)#port vlan 2 0/19 0
huawei(config)#interface vlanif 2
```

```
huawei(config-if-vlanif2)#ip address 8.1.2.1 24
huawei(config-if-vlanif2)#quit
```

2. Enable the BGP function.

```
huawei(config)#bgp 2000
huawei(config-bgp)#router-id 1.1.1.1
huawei(config-bgp)#peer 8.1.1.1 as-number 2000
huawei(config-bgp)#peer 8.1.2.2 as-number 2000
huawei(config-bgp)#quit
```

3. Save the data.

```
huawei(config)#save
```

----End

Result

- Run the **display bgp peer** command, and you can query the configuration.

Configuration File

```
vlan 6 smart
port vlan 6 0/19 0
interface vlanif 6
ip address 8.1.1.2 24
quit
vlan 2 smart
port vlan 2 0/19 0
interface vlanif 2
ip address 8.1.2.1 24
quit
bgp 2000
router-id 1.1.1.1
peer 8.1.1.1 as-number 2000
peer 8.1.2.2 as-number 2000
quit
save
```

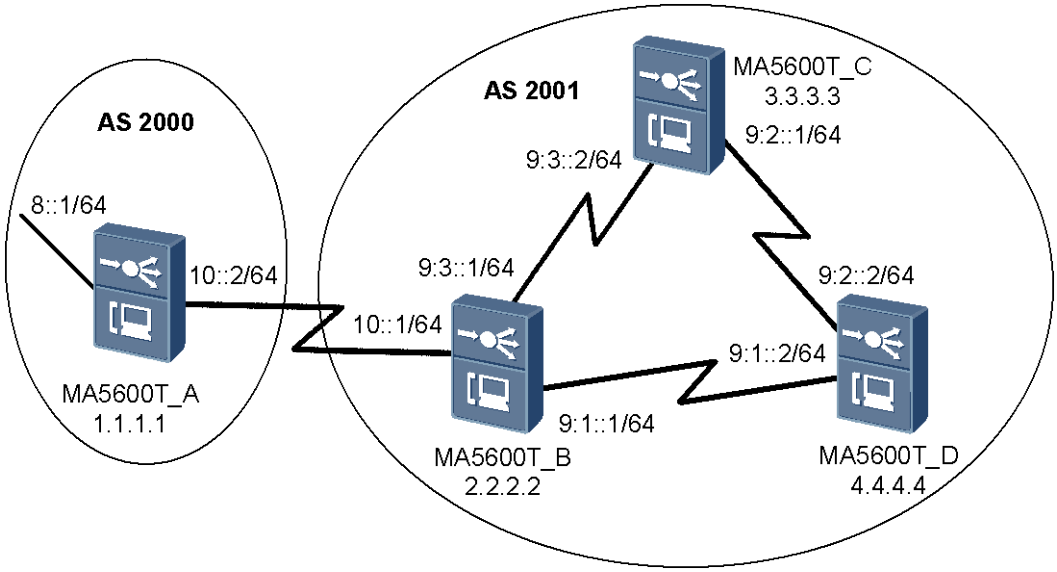
16.11.13 Configuration Example of BGP4+

This topic provides an example for configuring the BGP4+ on the device.

Service Requirements

In this example network, an external Border Gateway Protocol (EBGP) connection is set up between MA5600T_A and MA5600T_B, and an Interior Border Gateway Protocol (IBGP) connection is set up among MA5600T_B, MA5600T_C, and MA5600T_D.

Figure 16-85 Example network for configuring the BGP4+



Data Plan

Table 16-41 provides the data plan for configuring the BGP4+.

Table 16-41 Data plan for configuring the BGP4+

Item	Data	Remarks
MA5600T_A	IPv6 address of virtual local area network (VLAN) interface 6: 10::2/64	It is used for the EBGP connection to Autonomous System (AS) 2001.
	IPv6 address of VLAN interface 2: 8::1/64	-
	Router ID: 1.1.1.1	-
	AS number: 2000	-
MA5600T_B	IPv6 address of VLAN interface 6: 10::1/64	It is used for the EBGP connection to AS 2000.
	IPv6 address of VLAN interface 3: 9:3::1/64	It is used for the IBGP connection to the MA5600T_C.
	IPv6 address of VLAN interface 4: 9:1::1/64	It is used for the IBGP connection to the MA5600T_D.
	Router ID: 2.2.2.2	-
	AS number: 2001	-

Item	Data	Remarks
MA5600T_C	IPv6 address of VLAN interface 3: 9:3::2/64	It is used for the IBGP connection to the MA5600T_B.
	IPv6 address of VLAN interface 4: 9:2::1/64	It is used for the IBGP connection to the MA5600T_D.
	Router ID: 3.3.3.3	-
	AS number: 2001	-
MA5600T_D	IPv6 address of VLAN interface 5: 9:2::2/64	It is used for the IBGP connection to the MA5600T_C.
	IPv6 address of VLAN interface 4: 9:1::2/64	It is used for the IBGP connection to the MA5600T_B.
	Router ID: 4.4.4.4	-
	AS number: 2001	-

Procedure

Configure MA5600T_A.

1. Configure the IPv6 address of the Layer 3 interface.

```
huawei(config)#ipv6
huawei(config)#vlan 6 smart
huawei(config)#port vlan 6 0/19 0
huawei(config)#interface vlanif 6
huawei(config-if-vlanif6)#ipv6 enable
huawei(config-if-vlanif6)#ipv6 address 10::2 64
huawei(config-if-vlanif6)#quit
huawei(config)#vlan 2 smart
huawei(config)#port vlan 2 0/19 0
huawei(config)#interface vlanif 2
huawei(config-if-vlanif2)#ipv6 enable
huawei(config-if-vlanif2)#ipv6 address 8::1/64
huawei(config-if-vlanif2)#quit
```

2. Enable the Border Gateway Protocol (BGP) function and configure the EBGP neighbor between MA5600T_B.

```
huawei(config)#bgp 2000
huawei(config-bgp)#router-id 1.1.1.1
huawei(config-bgp)#peer 10::1 as-number 2001
huawei(config-bgp)#ipv6-family unicast
huawei(config-bgp-af-ipv6)#peer 10::1 enable
huawei(config-bgp-af-ipv6)#network 10:: 64
huawei(config-bgp-af-ipv6)#network 8:: 64
```

```
huawei(config-bgp-af-ipv6)#quit
huawei(config-bgp)#quit
```

3. Save the data.

```
huawei(config)#save
```

Step 1 Configure MA5600T_B.

1. Configure the IPv6 address of the Layer 3 interface.

```
huawei(config)#ipv6
huawei(config)#vlan 6 smart
huawei(config)#port vlan 6 0/19 0
huawei(config)#interface vlanif 6
huawei(config-if-vlanif6)#ipv6 enable
huawei(config-if-vlanif6)#ipv6 address 10::1 64
huawei(config-if-vlanif6)#quit
huawei(config)#vlan 3 smart
huawei(config)#port vlan 3 0/19 0
huawei(config)#interface vlanif 3
huawei(config-if-vlanif3)#ipv6 enable
huawei(config-if-vlanif3)#ipv6 address 9:3::1 64
huawei(config-if-vlanif3)#quit
huawei(config)#vlan 4 smart
huawei(config)#port vlan 4 0/19 0
huawei(config)#interface vlanif 4
huawei(config-if-vlanif4)#ipv6 enable
huawei(config-if-vlanif4)#ipv6 address 9:1::1 64
huawei(config-if-vlanif4)#quit
```

2. Enable the BGP function. Configure the EBGP neighbor between MA5600T_B and MA5600T_A, and the IBGP neighbor between MA5600T_B, MA5600T_C, and MA5600T_D.

```
huawei(config)#bgp 2001
huawei(config-bgp)#router-id 2.2.2.2
huawei(config-bgp)#peer 10::2 as-number 2000
huawei(config-bgp)#peer 9:3::2 as-number 2001
huawei(config-bgp)#peer 9:1::2 as-number 2001
huawei(config-bgp)#ipv6-family unicast
huawei(config-bgp-af-ipv6)#peer 10::2 enable
huawei(config-bgp-af-ipv6)#peer 9:3::2 enable
huawei(config-bgp-af-ipv6)#peer 9:1::2 enable
huawei(config-bgp-af-ipv6)#import-route direct
huawei(config-bgp-af-ipv6)#quit
huawei(config-bgp)#quit
```

3. Save the data.

```
huawei(config)#save
```

Step 2 Configure MA5600T_C.

1. Configure the IPv6 address of the Layer 3 interface.

```
huawei(config)#ipv6
huawei(config)#vlan 3 smart
huawei(config)#port vlan 3 0/19 0
huawei(config)#interface vlanif 3
huawei(config-if-vlanif3)#ipv6 enable
huawei(config-if-vlanif3)#ipv6 address 9:3::2 64
```

```
huawei(config-if-vlanif3)#quit
huawei(config)#vlan 5 smart
huawei(config)#port vlan 5 0/19 0
huawei(config)#interface vlanif 5
huawei(config-if-vlanif5)#ipv6 enable
huawei(config-if-vlanif5)#ipv6 address 9:2::1 64
huawei(config-if-vlanif5)#quit
```

2. Enable the BGP function. Configure the IBGP neighbor between MA5600T_B and MA5600T_D.

```
huawei(config)#bgp 2001
huawei(config-bgp)#router-id 3.3.3.3
huawei(config-bgp)#peer 9:3::1 as-number 2001
huawei(config-bgp)#peer 9:2::2 as-number 2001
huawei(config-bgp)#ipv6-family unicast
huawei(config-bgp-af-ipv6)#peer 9:3::1 enable
huawei(config-bgp-af-ipv6)#peer 9:2::2 enable
huawei(config-bgp-af-ipv6)#import-route direct
huawei(config-bgp-af-ipv6)#quit
huawei(config-bgp)#quit
```

3. Save the data.

```
huawei(config)#save
```

Step 3 Configure MA5600T_D.

1. Configure the IPv6 address of the Layer 3 interface.

```
huawei(config)#ipv6
huawei(config)#vlan 4 smart
huawei(config)#port vlan 4 0/19 0
huawei(config)#interface vlanif 4
huawei(config-if-vlanif4)#ipv6 enable
huawei(config-if-vlanif4)#ipv6 address 9:1::2 64
huawei(config-if-vlanif4)#quit
huawei(config)#vlan 5 smart
huawei(config)#port vlan 5 0/19 0
huawei(config)#interface vlanif 5
huawei(config-if-vlanif5)#ipv6 enable
huawei(config-if-vlanif5)#ipv6 address 9:2::2 64
huawei(config-if-vlanif5)#quit
```

2. Enable the BGP function. Configure the IBGP neighbor between MA5600T_B and MA5600T_C.

```
huawei(config)#bgp 2001
huawei(config-bgp)#router-id 4.4.4.4
huawei(config-bgp)#peer 9:1::2 as-number 2001
huawei(config-bgp)#peer 9:2::1 as-number 2001
huawei(config-bgp)#ipv6-family unicast
huawei(config-bgp-af-ipv6)#peer 9:1::2 enable
huawei(config-bgp-af-ipv6)#peer 9:2::1 enable
huawei(config-bgp-af-ipv6)#import-route direct
huawei(config-bgp-af-ipv6)#quit
huawei(config-bgp)#quit
```

3. Save the data.

```
huawei(config)#save
```

----End

Result

- Run the **display bgp peer** command, and you can see that:
  - The EBGP connection is set up between MA5600T_A and MA5600T_B.
  - The IBGP connections are set up among MA5600T_B, MA5600T_C, and MA5600T_D.
  - The route with the destination subnet 8::/64 exists on MA5600T_C and MA5600T_D, and the next hop of the route is the interface address of MA5600T_A
- Run the **ping ipv6** command on MA5600T_C and MA5600T_D to ping the Layer 3 interface (8::1/64) on MA5600T_A. The **ping ipv6** command is executed successfully.

Configuration File

Configuration on each MA5600T is similar. Take MA5600T_A for example.

```
ipv6
vlan 6 smart
port vlan 6 0/19 0
interface vlanif 6
ipv6 enable
ipv6 address 10::2 64
quit
vlan 2 smart
port vlan 2 0/19 0
interface vlanif 2
ipv6 enable
ipv6 address 8::1 64
quit
bgp 2000
router-id 1.1.1.1
peer 10::1 as-number 2001
ipv6-family unicast
network 8.0.0.0 8
peer 10::1 enable
network 10:: 64
network 8:: 64
quit
quit
```

16.11.14 References

The following table lists the references that apply in this chapter.

Table 16-42 References

Document No.	Document Name	Protocol Compliance
RFC 827	Exterior Gateway Protocol (EGP)	Fully compliant
RFC 1997	BGP Communities Attribute	Fully compliant

Document No.	Document Name	Protocol Compliance
RFC 2439	BGP Route Flap Damping	Fully compliant
RFC 2918	Route Refresh Capability for BGP-4	Fully compliant
RFC 3065	Autonomous System Confederations for BGP	Fully compliant
RFC 3232	Assigned Numbers: RFC 1700 is Replaced by an On-line Database	Fully compliant
RFC 3392	Capabilities Advertisement with BGP-4	Fully compliant
RFC 3682	The Generalized TTL Security Mechanism (GTSM)	Fully compliant
RFC 4271	A Border Gateway Protocol 4 (BGP-4)	Fully compliant
RFC 4456	BGP Route Reflection	Fully compliant
RFC 4486	Subcodes for BGP Cease Notification Message	Partially compliant: The backoff mechanism is not supported.
RFC 4724	Graceful Restart Mechanism for BGP	Fully compliant
RFC 4760	Multiprotocol Extensions for BGP-4	Fully compliant
draft-rijsman-bfd-down-subcode-00	BFD Down Subcode for BGP Cease Notification Message	Fully compliant
draft-ietf-idr-aigp-08	The Accumulated IGP Metric Attribute for BGP	Fully compliant

16.12 VRF

Virtual route forwarding instance (VRF) is a mechanism in which a device works as multiple virtual routing devices. After the Layer 3 interfaces of the device are divided into different VRFs, multiple route forwarding instances can be emulated on the device.

16.12.1 Introduction to VRF

Definition

VRF is an Layer 3 virtual private network (L3VPN). VRF is a mechanism in which a device works as multiple virtual routing devices. After the Layer 3 interfaces of the device are

divided into different VRFs, multiple route forwarding instances can be emulated on the device.

Purpose

Multiple virtual routing devices can be created on the MA5600T/MA5603T/MA5608T. That is, multiple L3VPNs can be established to implement the Layer 3 isolation and independent packet forwarding among different VRFs. Moreover, in different VRFs, the IP address can be reused, and also DHCP relay multi-instances, routing multi-instances, and independent route forwarding tables are supported.

The MA5600T/MA5603T/MA5608T categorizes VRFs by VLANs to provide L3VPN solutions. All the packets or related protocols on the Layer 3 interface of a VRF are processed only in this VRF, which is unrelated to other VRFs. In this way, the services or users can be isolated, and the IP addresses can be saved.

VRF has two application scenarios:

- When the triple play service is provisioned to xDSL access users or GPON access users, different services are isolated from each other by VRF, and all services of the device are carried and go upstream by the same physical link. One VLAN Layer 3 interface can be bound to only one VR, and the upstream port belongs to multiple Layer 3 interfaces. Different VLAN Layer 3 interfaces are bound to different VRs, and each VR forwards data according to the route learned by this VR.
- When the triple play service is provisioned to xDSL access users or GPON access users, different services are isolated from each other by VRF, and all services of the device are carried and go upstream by two or more physical links. The links in this case are in the Layer 3 mode, and different services are isolated from each other by VRF.

The difference of the two scenarios is that dual or multiple links are adopted for upstream transmission in scenario 2, where the effect of different VRs going different "ways" is more vivid.

16.12.2 VRF Principle

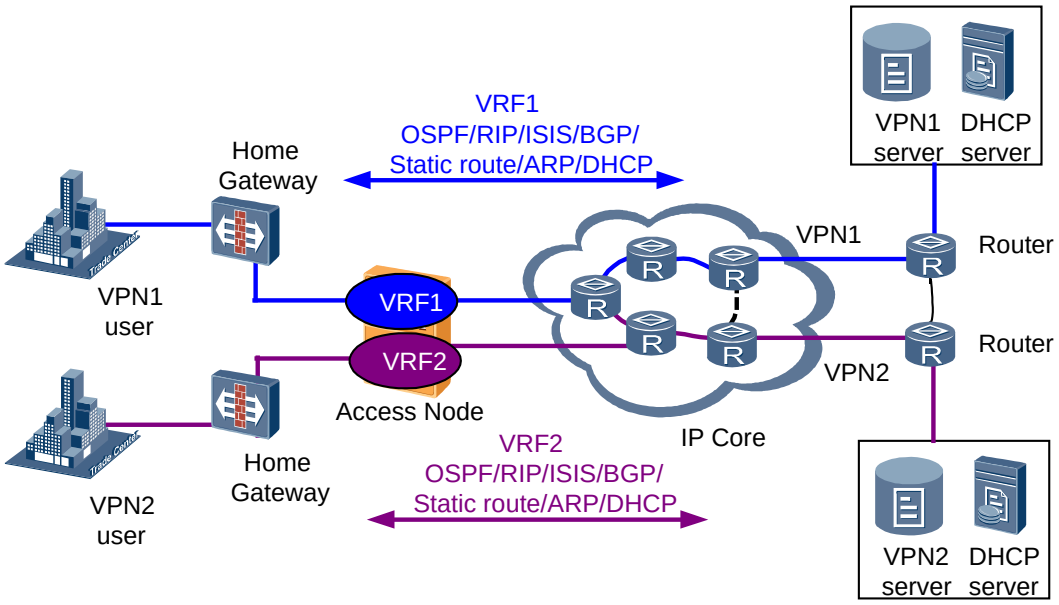
VRF Compatibility

The VRF architecture is compatible with the virtual private routed network (VPRN) architecture as defined in RFC2764.

VRF Architecture

VRF is an architecture of IP networks, as shown in Figure 16-86. When users are isolated by service types or ISPs, or the users of different VPNs are prohibited from communicating with each other, multiple L3VPNs must be established in an IP network.

Figure 16-86 VRF architecture



The MA5600T/MA5603T/MA5608T VRF supports the following functions:

- Creating a VRF instance

  You can create a VRF instance and set the name of the VRF as the reference flag through the CLI.
- Adding the VLAN Layer 3 interface and the loopback interface into a VRF instance
  - The MA5600T/MA5603T/MA5608T differentiates VRFs by the VLAN Layer 3 interfaces. A VRF contains one or more VLAN Layer 3 interfaces. When receiving or transmitting packets, any VLAN Layer 3 interface that belongs to the VRF must use the Layer 3 route forwarding table of the VRF. Moreover, the packets in the VRF must be forwarded between these VLAN Layer 3 interfaces and cannot be forwarded to any other VLAN Layer 3 interfaces that do not belong to the VRF.
  - After a loopback interface is bound with the VRF instance, the loopback interface can process all the routing protocols in the VRF.
  - The IP addresses configured in the VLAN Layer 3 interfaces of different VRFs can be identical, but the IP addresses in the same VRF cannot be identical.
- Isolating ARP in a VRF

  The ARP in different VRFs is isolated, but the user IP addresses in different VRFs can be identical.
- Supporting independent ISIS, OSPF, RIP, or BGP routing protocol process for different VRFs
- Supporting the Layer 3 DHCP relay or DHCP proxy in a VRF

  The MA5600T/MA5603T/MA5608T supports the DHCP configuration based on the VLAN to implement the DHCP relay or DHCP proxy function in the VRF.
- Supporting the ping and trace route functions in a VRF
  - Ping and trace route are the basic network maintenance means.
  - The ping function is used to check the connectivity and reachability of a remote host by sending the ping packets to the host.

- The trace route function is used to check the network connectivity and locate the network faults by testing the route that the data packets pass through from the host to the destination.

16.12.3 Configuring IPv4 in VPN

This topic describes how to categorize virtual private network (VPN) instances by VLANs, and realize the virtual IPv4 static route forwarding in different VPN instances.

Context

- A VPN instance is also called a VPN Routing and Forwarding (VRF) table. VRF is a Layer 3 virtual private network (L3 VPN). VRF is a mechanism in which a device works as multiple virtual routing devices. After the Layer 3 interfaces of the device are divided into different VRFs, multiple route forwarding instances can be emulated on the device.
- Multiple virtual routing devices can be created on the access node. That is, multiple L3VPNs can be established to implement the Layer 3 isolation and independent packet forwarding among different VRFs. The access node supports the following VRF functions:
  - In different VRF instances, the IP address can be reused. It means that the IP addresses of the Layer 3 interfaces which belong to different VRF instances can be the same.
  - The ping and trace route functions are supported in a VRF.
  - The users of different VRF instances can obtain the IP addresses through the Dynamic Host Control Protocol (DHCP) relay or the DHCP proxy.
  - The static routes and the dynamic routes in a VRF instance do not affect each other, and the routing entry in each VRF instance supports the routing function independently.

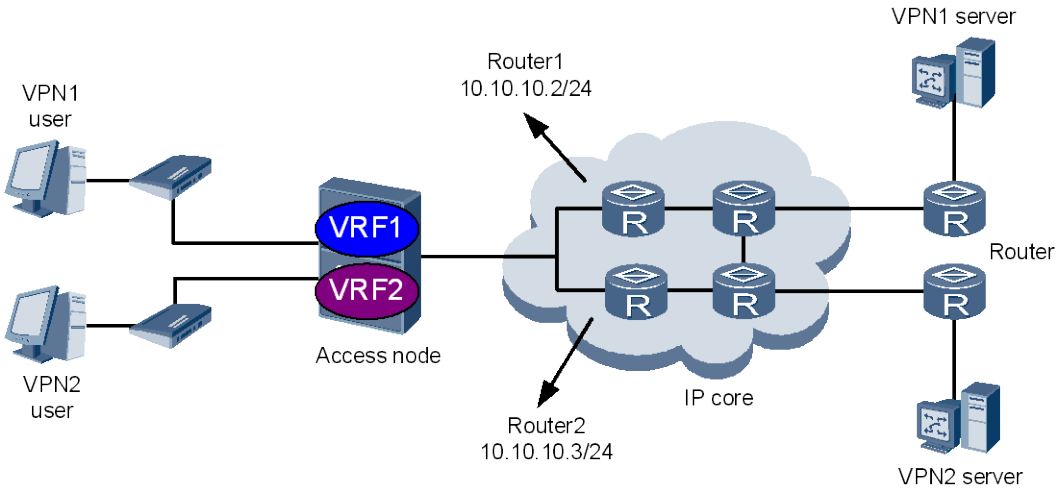
Networking

Figure 16-87 shows an example network for configuring IPv4 in VPN.

The access node categorizes VRF instances by VLANs to provide L3VPN solutions. In this example, VPN instance VRF1 is categorized by virtual local area network (VLAN) 200, and IPv4 static routes are added in the virtual route forwarding entries of VRF1. The access node selects the routes for the users of VPN1 by querying the routing entries of VRF1. Similarly, VPN instance VRF2 is categorized by VLAN 300 and is used to select the routes for the users of VPN2. The access node implements the Layer 3 isolation and independent packet forwarding through different VRF instances.

This example describes how to configure the function of virtual static route forwarding by adding IPv4 static routes application on the instance. The function of virtual dynamic route forwarding can be realized by enabling the process of the dynamic routing protocols such as the open shortest path first (OSPF), Routing Information Protocol (RIP), intermediate system to intermediate system (IS-IS), and Border Gateway Protocol (BGP) in a VRF instance.

Figure 16-87 Example network for configuring IPv4 in VPN



Data Plan

Table 16-43 provides the data plan for configuring IPv4 in VPN.

Table 16-43 Data plan for configuring IPv4 in VPN

Item	Data
VRF1 (for VPN1)	Name of the VPN instance: vpn1 Route distinguisher (RD) of the VPN instance: 100:1
	Upstream port: 0/19/0 VLAN: 200 VLAN type: smart VLAN VPN1 user: • Gigabit-capable passive optical network (GPON) port: 0/2/0 • ONT ID: 0 • GEM Port ID: 0
	IP address of the Layer 3 interface of VLAN 200: 10.10.10.1/24 IP address of router1: 10.10.10.2/24 IP address of the VPN1 server: 10.10.20.1/24
VRF2 (for VPN2)	Name of the VPN instance: vpn2 RD of the VPN instance: 100:2
	Upstream port: 0/19/0 VLAN: 300 VLAN type: smart VLAN VPN2 user:

Item	Data
	• GPON port: 0/2/1 • ONT ID: 1 • GEM Port ID: 1
	IP address of the Layer 3 interface of VLAN 300: 10.10.10.1/24 IP address of router2: 10.10.10.3/24 IP address of the VPN2 server: 10.10.30.1/24

Procedure

- Configure VRF1 (for VPN1).
  a. Create a VPN instance of the IPv4 address family.

```
huawei(config)#ip vpn-instance vpn1
huawei(config-vpn-instance-vpn1)#ipv4-family
```

  b. Configure the RD of the VPN instance.

```
huawei(config-vpn-instance-vpn1-af-ipv4)#route-distinguisher 100:1
huawei(config-vpn-instance-vpn1-af-ipv4)#quit
huawei(config-vpn-instance-vpn1)#quit
```

  c. Create a smart VLAN and add the upstream port and the service port to it.

```
huawei(config)#vlan 200 smart
huawei(config)#port vlan 200 0/19 0
huawei(config)#service-port vlan 200 gpon 0/2/0 ont 0 gemport 0 multi-service user-8021p 0 user-vlan 200 rx-cttr 5 tx-cttr 5
```

  d. Associate the Layer 3 interface with the VPN instance.

```
huawei(config)#interface vlanif 200
huawei(config-if-vlanif200)#ip binding vpn-instance vpn1
  Info: All IPv4 related configurations on this interface are removed
  Info: All IPv6 related configurations on this interface are removed
```

  e. Configure the IP address of the VLAN Layer 3 interface.

```
huawei(config-if-vlanif200)#ip address 10.10.10.1 24
huawei(config-if-vlanif200)#quit
```

  f. Configure the IPv4 static route.

```
huawei(config)#ip route-static vpn-instance vpn1 10.10.20.0 24 10.10.10.2
```

  g. Save the data.

```
huawei(config)#save
```

- Configure VRF2 (for VPN2).
  a. Create a VPN instance of the IPv4 address family.

```
huawei(config)#ip vpn-instance vpn2
huawei(config-vpn-instance-vpn2)#ipv4-family
```

  b. Configure the RD of the VPN instance.

```
huawei(config-vpn-instance-vpn1-af-ipv4)#route-distinguisher 100:2
huawei(config-vpn-instance-vpn1-af-ipv4)#quit
huawei(config-vpn-instance-vpn1)#quit
```

c. Create a smart VLAN and add the upstream port and the service port to it.

```
huawei(config)#vlan 300 smart
huawei(config)#port vlan 300 0/19 0
huawei(config)#service-port vlan 300 gpon 0/2/1 ont 1 gemport 1 multi-service
user-8021p 0 user-vlan 300 rx-cttr 6 tx-cttr 6
```

d. Associate the Layer 3 interface with the VPN instance.

```
huawei(config)#interface vlanif 300
huawei(config-if-vlanif300)#ip binding vpn-instance vpn2
  Info: All IPv4 related configurations on this interface are removed
  Info: All IPv6 related configurations on this interface are removed
```

e. Configure the IP address of the VLAN Layer 3 interface.

```
huawei(config-if-vlanif300)#ip address 10.10.10.1 24
huawei(config-if-vlanif300)#quit
```

f. Configure the IPv4 static route.

```
huawei(config)#ip route-static vpn-instance vpn2 10.10.30.0 24 10.10.10.3
```

g. Save the data.

```
huawei(config)#save
```

----End

Result

Run the **display ip vpn-instance** command to query the VPN configuration.

```
huawei(config)#display ip vpn-instance
{ <cr>|import-vt<K>|interface<K>|STRING<1-31>|verbose<K>||<K> }:

  Command:
        display ip vpn-instance
 Total VPN-Instances configured : 2

  VPN-Instance Name            Address-family
  vpn1                        ipv4
  vpn2                        ipv4
```

Run the following commands to verify that the VRF instances are configured successfully. The two IPv4 static routes are added to the IP routing table of VPN1 and VPN2.

```
huawei(config)#display ip routing-table vpn-instance vpn1
{ <cr>|verbose<K>|statistics<K>|protocol<K>|acl<K>|ip-prefix<K>|ip_addr<I><X.X.X
.X> }:

  Command:
        display ip routing-table vpn-instance vpn1
Routing Tables: vpn1
       Destinations : 3       Routes : 3
```

```
Destination/Mask    Proto  Pre  Cost      NextHop         Interface

    10.10.10.0/24  Direct 0    0         10.10.10.1      vlanif200
    10.10.10.1/32  Direct 0    0         127.0.0.1       InLoopBack0
    10.10.20.0/24  Static 60   0         10.10.10.2      vlanif200

huawei(config)#display ip routing-table vpn-instance vpn2
{ <cr>|verbose<K>|statistics<K>|protocol<K>|acl<K>|ip-prefix<K>|ip_addr<I><X.X.X
.X> }:

  Command:
        display ip routing-table vpn-instance vpn2
Routing Tables: vpn2
       Destinations : 3       Routes : 3

Destination/Mask    Proto  Pre  Cost      NextHop         Interface

    10.10.10.0/24  Direct 0    0         10.10.10.1      vlanif300
    10.10.10.1/32  Direct 0    0         127.0.0.1       InLoopBack0
    10.10.30.0/24  Static 60   0         10.10.10.3      vlanif300
```

Run the **ping** and **tracert** commands to check the VPN connectivity.

The MA5600T/MA5603T/MA5608T categorizes VRF instances by VLANs to provide L3VPN solutions, realizing the Layer 3 isolation of users or services.

- For the users of VPN1, the MA5600T/MA5603T/MA5608T selects the routes by querying the routing entries of VPN1. For example, for the packets to be sent to the VPN1 server (with IP address 10.10.20.1), the MA5600T/MA5603T/MA5608T selects its next hop router (with IP address 10.10.10.2) to forward the packets.
- For the users of VPN2, the MA5600T/MA5603T/MA5608T selects the routes by querying the routing entries of VPN2. For example, for the packets to be sent to the VPN2 server (with IP address 10.10.30.1), the MA5600T/MA5603T/MA5608T selects its next hop router (with IP address 10.10.10.3) to forward the packets.
- For the users outside the VPNs, the route to the VPN1 server or the VPN2 server is not available.

Configuration File

Only the configuration files related to the VPN are listed.

```
ip vpn-instance vpn1
ipv4-family
route-distinguisher 100:1
quit
quit
ip vpn-instance vpn2
ipv4-family
route-distinguisher 100:2
quit
quit
interface vlanif200
ip binding vpn-instance vpn1
ip address 10.10.10.1/24
quit
interface vlanif300
```

```
ip binding vpn-instance vpn2
ip address 10.10.10.1/24
quit
ip route-static vpn-instance vpn1 10.10.20.0 24 10.10.10.2
ip route-static vpn-instance vpn2 10.10.30.0 24 10.10.10.3
```

16.12.4 Configuring IPv6 in VPN

This topic describes how to categorize virtual private network (VPN) instances by virtual local area networks (VLANs), and implement the virtual IPv6 static route forwarding in different VPN instances.

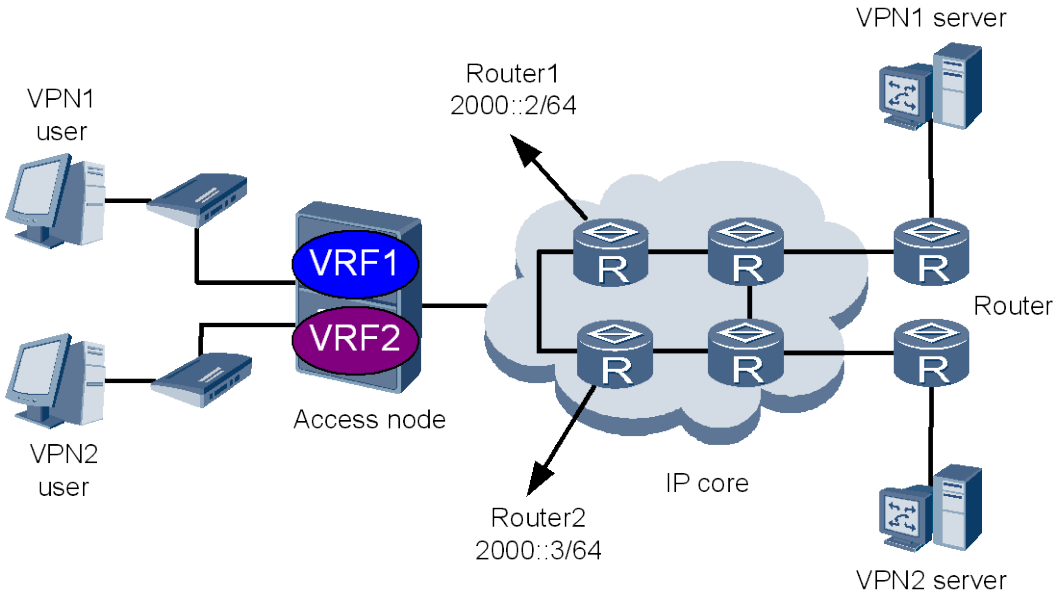
Networking

The access node supports virtual route forwarding (VRF) in the IPv6 network. The VRF principle and functions in the IPv6 network are the same as those in the IPv4 network.Figure 16-88 shows an example network for configuring IPv6 in VPN.

The access node categorizes VRF instances by VLANs to provide L3 VPN solutions. In this example, VPN instance VRF1 is categorized by virtual local area network (VLAN) 200, and IPv6 static routes are added in the virtual route forwarding entries of VRF1. The access node selects the routes for the users of VPN1 by querying the routing entries of VRF1. Similarly, VPN instance VRF2 is categorized by VLAN 300 and is used to select the routes for the users of VPN2. The access node implements the Layer 3 isolation and independent packet forwarding through different VRF instances.

This example describes how to configure the function of virtual static route forwarding by adding IPv6 static routes application on the instance. In addition, virtual dynamic route forwarding can be implemented by using BGP4+. For details about BGP4+ configurations, see 16.11.13 Configuration Example of BGP4+

Figure 16-88 Example network for configuring IPv6 in VPN



Data Plan

Table 16-44 provides the data plan for configuring IPv6 in VPN.

Table 16-44 Data plan for configuring IPv6 in VPN

Item	Data
VRF1 (for VPN1)	Name of the VPN instance: vpn1 Route distinguisher (RD) of the VPN instance: 100:1
	Upstream port: 0/19/0 VLAN: 200 VLAN type: smart VLAN VPN1 user: • Gigabit-capable passive optical network (GPON) port: 0/2/0 • ONT ID: 0 • GEM Port ID: 0
	IPv6 address of the Layer 3 interface of VLAN 200: 2000::1/64 IPv6 address of router1: 2000::2/64 IPv6 address of the VPN1 server: 2001::1/64
VRF2 (for VPN2)	Name of the VPN instance: vpn2 Route distinguisher (RD) of the VPN instance: 100:2
	Upstream port: 0/19/0 VLAN: 300 VLAN type: smart VLAN VPN2 user: • GPON port: 0/2/1 • ONT ID: 1 • GEM Port ID: 1
	IPv6 address of the Layer 3 interface of VLAN 300: 2000::1/64 IPv6 address of router2: 2000::3/64 IPv6 address of the VPN2 server: 2002::1/64

Procedure

- Configure VRF1 (for VPN1).
  a. Enable IPv6.

```
huawei(config)#ipv6
```

  b. Create a VPN instance of the IPv6 address family.

```
huawei(config)#ip vpn-instance vpn1
huawei(config-vpn-instance-vpn1)#ipv6-family
```

c. Configure the RD of the VPN instance.

```
huawei(config-vpn-instance-vpn1-af-ipv6)#route-distinguisher 100:1
```

d. (Optional) Configure the IPv6 routing specifications of the VPN instance.

```
huawei(config-vpn-instance-vpn1-af-ipv6)#prefix limit 1000 simply-alert
huawei(config-vpn-instance-vpn1-af-ipv6)#routing-table limit 1000
simply-alert
huawei(config-vpn-instance-vpn1-af-ipv6)#quit
huawei(config-vpn-instance-vpn1)#quit
```

e. Create a smart VLAN and add the upstream port and the service port to it.

```
huawei(config)#vlan 200 smart
huawei(config)#port vlan 200 0/19 0
huawei(config)#service-port vlan 200 gpon 0/2/0 ont 0 gemport 0 multi-service
user-8021p 0 user-vlan 200 rx-cttr 5 tx-cttr 5
```

f. Associate the Layer 3 interface with the VPN instance.

```
huawei(config)#interface vlanif 200
huawei(config-if-vlanif200)#ip binding vpn-instance vpn1
  Info: All IPv4 related configurations on this interface are removed
  Info: All IPv6 related configurations on this interface are removed
```

g. Configure the IPv6 address of the VLAN Layer 3 interface.

```
huawei(config-if-vlanif200)#ipv6 enable
huawei(config-if-vlanif200)#ipv6 address 2000::1 64
huawei(config-if-vlanif200)#quit
```

h. Configure the IPv6 static route.

```
huawei(config)#ipv6 route-static vpn-instance vpn1 2001:: 64 2000::2
```

i. Save the data.

```
huawei(config)#save
```

- Configure VRF2 (for VPN2).

a. Enable IPv6.

```
huawei(config)#ipv6
```

b. Create a VPN instance of the IPv6 address family.

```
huawei(config)#ip vpn-instance vpn2
huawei(config-vpn-instance-vpn2)#ipv6-family
```

c. Configure the RD of the VPN instance.

```
huawei(config-vpn-instance-vpn2-af-ipv6)#route-distinguisher 100:2
```

d. (Optional) Configure the IPv6 routing specifications of the VPN instance.

```
huawei(config-vpn-instance-vpn2-af-ipv6)#prefix limit 1000 simply-alert
huawei(config-vpn-instance-vpn2-af-ipv6)#routing-table limit 1000
simply-alert
huawei(config-vpn-instance-vpn2-af-ipv6)#quit
huawei(config-vpn-instance-vpn2)#quit
```

e. Create a smart VLAN and add the upstream port and the service port to it.

```
huawei(config)#vlan 300 smart
huawei(config)#port vlan 300 0/19 0
huawei(config)#service-port vlan 300 gpon 0/2/1 ont 1 gemport 1 multi-service
user-8021p 0 user-vlan 300 rx-cttr 6 tx-cttr 6
```

f. Associate the Layer 3 interface with the VPN instance.

```
huawei(config)#interface vlanif 300
huawei(config-if-vlanif300)#ip binding vpn-instance vpn2
  Info: All IPv4 related configurations on this interface are removed
  Info: All IPv6 related configurations on this interface are removed
```

g. Configure the IPv6 address of the VLAN Layer 3 interface.

```
huawei(config-if-vlanif300)#ipv6 enable
huawei(config-if-vlanif300)#ipv6 address 2000::1 64
huawei(config-if-vlanif300)#quit
```

h. Configure the IPv6 static route.

```
huawei(config)#ipv6 route-static vpn-instance vpn2 2002:: 64 2000::3
```

i. Save the data.

```
huawei(config)#save
```

----End

Result

Run the **display ip vpn-instance** command to query the VPN configurations.

```
huawei(config)#display ip vpn-instance
{ <cr>|import-vt<K>|interface<K>|STRING<1-31>|verbose<K>||<K> }:

  Command:
        display ip vpn-instance
 Total VPN-Instances configured : 2

  VPN-Instance Name            Address-family
  vpn1                        ipv6
  vpn2                        ipv6
```

Run the following commands to verify that the configurations are successful and the IPv6 static route is added to the IPv6 routing table of VPN1 and VPN2.

```
huawei(config)#display ipv6 routing-table vpn-instance vpn1
{ <cr>|acl<K>|ipv6-prefix<K>|protocol<K>|statistics<K>|verbose<K>|x:x::x:x<IPv6>
<x:x::x:x>||<K> }:

  Command:
        display ipv6 routing-table vpn-instance vpn1
Routing Table : vpn1
       Destinations : 4       Routes : 4

 Destination  : 2000::                     PrefixLength : 64
 NextHop     : 2000::1                    Preference   : 0
 Cost        : 0                         Protocol    : Direct
 RelayNextHop : ::                         TunnelID    : 0x0
```

```
 Interface    : vlanif200                   Flags       : D

 Destination  : 2000::1                     PrefixLength : 128
 NextHop      : ::1                        Preference   : 0
 Cost         : 0                          Protocol     : Direct
 RelayNextHop : ::                          TunnelID     : 0x0
 Interface    : InLoopBack0                 Flags        : D

 Destination  : 2001::                      PrefixLength : 64
 NextHop      : 2000::2                    Preference   : 60
 Cost         : 0                          Protocol     : Static
 RelayNextHop : ::                          TunnelID     : 0x0
 Interface    : vlanif200                   Flags        : RD

 Destination  : FE80::                      PrefixLength : 10
 NextHop      : ::                         Preference   : 0
 Cost         : 0                          Protocol     : Direct
 RelayNextHop : ::                          TunnelID     : 0x0
 Interface    : null0                       Flags        : D
huawei(config)#display ipv6 routing-table vpn-instance vpn2
{ <cr>|acl<K>|ipv6-prefix<K>|protocol<K>|statistics<K>|verbose<K>|x:x::x:x<IPv6>
<x:x::x:x>||<K> }:

  Command:
        display ipv6 routing-table vpn-instance vpn2
Routing Table : vpn2
       Destinations : 4       Routes : 4

 Destination  : 2000::                      PrefixLength : 64
 NextHop      : 2000::1                    Preference   : 0
 Cost         : 0                          Protocol     : Direct
 RelayNextHop : ::                          TunnelID     : 0x0
 Interface    : vlanif200                   Flags        : D

 Destination  : 2000::1                     PrefixLength : 128
 NextHop      : ::1                        Preference   : 0
 Cost         : 0                          Protocol     : Direct
 RelayNextHop : ::                          TunnelID     : 0x0
 Interface    : InLoopBack0                 Flags        : D

 Destination  : 2002::                      PrefixLength : 64
 NextHop      : 2000::3                    Preference   : 60
 Cost         : 0                          Protocol     : Static
 RelayNextHop : ::                          TunnelID     : 0x0
 Interface    : vlanif300                   Flags        : RD

 Destination  : FE80::                      PrefixLength : 10
 NextHop      : ::                         Preference   : 0
 Cost         : 0                          Protocol     : Direct
 RelayNextHop : ::                          TunnelID     : 0x0
 Interface    : null0                       Flags        : D
```

Run the **ping ipv6** and **tracert ipv6** commands to check the VPN connectivity.

The MA5600T/MA5603T/MA5608T categorizes VRF instances by VLANs to provide L3 VPN solutions, realizing the Layer 3 isolation of users or services.

- For the users of VPN1, the MA5600T/MA5603T/MA5608T selects the routes by querying the routing entries of VPN1. For example, for the packets to be sent to the VPN1 server (with IPv6 address 2001::1), the MA5600T/MA5603T/MA5608T selects its next hop router (with IPv6 address 2000::2) to forward the packets.
- For the users of VPN2, the MA5600T/MA5603T/MA5608T selects the routes by querying the routing entries of VPN2. For example, for the packets to be sent to the VPN2 server (with IPv6 address 2002::1), the MA5600T/MA5603T/MA5608T selects its next hop router (with IPv6 address 2000::3) to forward the packets.
- For the users outside the VPNs, the route to the VPN1 server or the VPN2 server is not available.

Configuration File

Only the configuration files related to the VPN are listed.

```
ipv6
ip vpn-instance vpn1
ipv6-family
route-distinguisher 100:1
routing-table limit 1000 simply-alert
prefix limit 1000 simply-alert
quit
quit
ip vpn-instance vpn2
ipv6-family
route-distinguisher 100:2
routing-table limit 1000 simply-alert
prefix limit 1000 simply-alert
quit
quit
interface vlanif200
ip binding vpn-instance vpn1
ipv6 enable
ipv6 address 2000::1/64
quit
interface vlanif300
ip binding vpn-instance vpn2
ipv6 enable
ipv6 address 2000::1/64
quit
ipv6 route-static vpn-instance vpn1 2001:: 64 2000::2
ipv6 route-static vpn-instance vpn2 2002:: 64 2000::3
```

16.13 Routing policy

16.13.1 Introduction to Routing Policies

Definition

Routing policies are used to filter routes and set attributes for routes. Changing route attributes (including reachability) changes the path that network traffic passes through.

NOTE

The difference between a routing policy and policy-based routing (PBR) is as follows:

- Routing policies apply to routes. Based on routing protocols, the result of route generation, advertisement, and selection is changed by following rules, changing parameters, or using control modes. That is, the contents in the routing table are changed.
- PBR applies to data packets. PBR provides a means to route or forward data packets flexibly based on predefined policies instead of following the routes in the existing routing table.

Purpose

When advertising, receiving, and importing routes, the Router implements certain policies based on actual networking requirements to filter routes and change the attributes of the routes. Routing policies serve the following purposes:

- Control route advertising

  Only routes that match the rules specified in a policy are advertised.
- Control route receiving

  Only the required and valid routes are received. This reduces the size of the routing table and improves network security.
- Filter and control imported routes

  A routing protocol may import routes discovered by other routing protocols. Only routes that satisfy certain conditions are imported to meet the requirements of the protocol.
- Modify attributes of specified routes

  Attributes of the routes that are filtered by a routing policy are modified to meet the requirements of the local device.

Benefits

This feature brings the following benefits:

- Controls the size of the routing table, saving system resources.
- Controls route receiving and advertising, improving network security.
- Modifies attributes of routes for proper traffic planning, improving network performance.

16.13.2 References

None.

16.13.3 Principles

Implementation

Routing policies are implemented using the following procedures:

- Define rules: Define features of routes to which routing policies are applied. Users define a set of matching rules based on different attributes of routes, such as the destination address and the address of the router that advertises the routes.
- Implement the rules: Apply the matching rules to routing policies for advertising, receiving, and importing routes.

Filter

A filter is the core of a routing policy and is defined using a set of matching rules. The MA5600T/MA5603T/MA5608T provides the filters listed in Table 16-45.

Table 16-45 Comparisons between filters

Filter	Applicable Scope	Matching Rules
Access control list (ACL)	Dynamic routing protocols	Inbound interface, source or destination IP address, protocol type, and source or destination port number
IP prefix list	Dynamic routing protocols	Source and destination IP addresses and next hop address
AS_Path filter	BGP	AS_Path attribute
Community filter	BGP	Community attribute
Route-Policy	Dynamic routing protocols	Destination IP address, next hop address, cost, interface information, route type, ACL, IP prefix list, AS_Path filter, and community filter

The ACL, IP prefix list, AS_Path filter, and community filter can be used only to filter routes but not modify attributes of the filtered routes. A Route-Policy is a comprehensive filter, and it can use the matching rules of the ACL, IP prefix list, AS_Path filter, and community filter to filter routes. In addition, attributes of the filtered routes can be modified using the Route-Policy. The following section describes the filters in more detail.

ACL

An ACL is a set of sequential filtering rules. Users can define rules based on packet information, such as inbound interfaces, source or destination IP addresses, protocol types, or source or destination port numbers and specify an action to deny or permit packets. After an ACL is configured, the system classifies received packets based on the rules defined in the ACL and denies or permits the packets accordingly.

An ACL only classifies packets based on defined rules and can be used to filter packets only when it is applied to a routing policy.

ACLs can be configured for both IPv4 packets and IPv6 packets. Based on the usage, ACLs are classified into three types: interface-based ACLs, basic ACLs, and advanced ACLs. Users can specify the IP address and subnet address range in an ACL to match the source IP address, destination network segment address, or the next hop address of a route.

ACLs can be configured on access or core devices to:

- Protect the devices against IP, TCP, and ICMP packet attacks.
- Control network access. For example, ACLs can be used to control the access of enterprise network users to external networks, the specific network resources that users can access, and the period for which users can access networks.

- Limit network traffic and improve network performance. For example, ACLs can be used to limit bandwidth for upstream and downstream traffic, charge for the bandwidth that users have applied for, and fully use high-bandwidth network resources.

IP Prefix List

An IP prefix list contains a group of route filtering rules. Users can specify the prefix and mask length range to match the destination network segment address or the next hop address of a route. An IP prefix list is used to filter routes that are advertised and received by various dynamic routing protocols.

An IP prefix list is easier and more flexible than an ACL. However, if a large number of routes with different prefixes need to be filtered, configuring an IP prefix list to filter the routes is complex.

IP prefix lists can be configured for both IPv4 routes and IPv6 routes, and they share the same implementation process. An IP prefix list filters routes based on the mask length or mask length range.

- Mask length: An IP prefix list filters routes based on IP address prefixes. An IP address prefix is defined by an IP address and the mask length. For example, in a route to 10.1.1.1/16, the mask length is 16 bits, and the valid prefix is 16 bits (10.1.0.0).
- Mask length range: Routes with the IP address prefix and mask length within the range defined in the IP prefix list meet the matching rules.

NOTE

0.0.0.0 is a wildcard address. If the IP prefix is 0.0.0.0, users must specify either a mask or a mask length range, with the following results:

- If a mask is specified, all routes with the mask are permitted or denied as required.
- If a mask length range is specified, all routes with the mask length in the range are permitted or denied as required.

AS_Path Filter

An AS_Path filter is used to filter BGP routes based on AS_Path attributes contained in the BGP routes. The AS_Path attribute is used to record numbers of all ASs that a BGP route passes through from the local end to the destination in the distance-vector (DV) order. Therefore, filtering rules defined based on AS_Path attributes can be used to filter BGP routes.

The matching condition of an AS_Path filter is specified using a regular expression. For example, ^30 indicates that only the AS_Path attribute starting with 30 is matched. Using a regular expression can simplify configurations. For details about regular expressions, see Commands>CLI Operation Characteristics>Parameter.

NOTE

The AS_Path attribute is a private attribute of BGP and is therefore used to filter BGP routes only. For details about the AS_Path attribute, see 16.11.2 Basic Principle of BGP.

Community Filter

A community filter is used to filter BGP routes based on the community attributes contained in the BGP routes. The community attribute is a set of destination addresses with the same characteristics. Therefore, filtering rules defined based on community attributes can be used to filter BGP routes.

In addition to the well-known community attributes, users can define community attributes using digits. The matching condition of a community filter can be specified using a community ID or a regular expression.

NOTE

Like AS_Path filters, community filters are used to filter only BGP routes because the community attribute is also a private attribute of BGP. For details about the community attribute, see 16.11.6 Community Attribute.

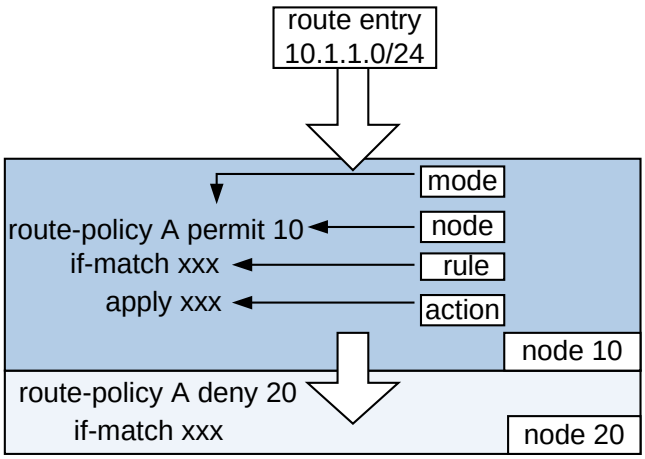
Route-Policy

A Route-Policy is a complex filter. It is used to match attributes of specified routes and change route attributes when specific conditions are met. A Route-Policy can use the preceding six filters to define its matching rules.

- **Composition of a Route-Policy**

  As shown in Figure 16-89, a Route-Policy consists of node IDs, matching mode, **if-match** clauses, and **apply** clauses.

  Figure 16-89 Composition of a Route-Policy



  - Node ID

    A Route-Policy consists of one or more nodes. Node IDs are specified as indexes in the IP prefix list. In a Route-Policy, routes are filtered based on the following rules:

    - Sequential matching: The system checks entries based on node IDs in ascending order. Therefore, specifying the node IDs in the required sequence is recommended.
    - One-time matching: The relationship between the nodes of a Route-Policy is "OR". If a route matches one node, the route matches the Route-Policy and will not be matched against the next node.

  - Matching mode

    Either of the following matching modes can be used:

    - **permit**: specifies the **permit** mode of a node. If a route matches the **if-match** clauses of a node, all the actions defined by **apply** clauses are performed, and the matching is complete. If a route does not match the **if-match** clauses of the node, the route continues to match the next node.

- **deny**: specifies the **deny** mode of a node. In the **deny** mode, the **apply** clauses are not used. If a route matches all the **if-match** clauses of the node, the route is denied by the node and the next node is not matched. If the entry does not match all the **if-match** clauses, the next node is matched.

NOTE

To allow other routes to pass through, a Route-Policy that contains no **if-match** or **apply** clause in the **permit** mode needs to be configured for a node next to multiple nodes that are in the **deny** mode.

- **if-match** clause

  The **if-match** clause defines the matching rules.

  Each node of a Route-Policy can comprise multiple **if-match** clauses or no **if-match** clause at all. If no **if-match** clause is configured for a node in the **permit** mode, all routes match the node.

- **apply** clause

  The **apply** clauses specify actions. When a route matches a Route-Policy, the system sets some attributes for the route based on the **apply** clause.

  Each node of a Route-Policy can comprise multiple **apply** clauses or no **apply** clause at all. The **apply** clause is not used when routes need to be filtered but attributes of the routes do not need to be set.

- **Matching results of a Route-Policy**

  The matching results of a Route-Policy are obtained based on the following aspects:

  - Matching mode of the node, either **permit** or **deny**
  - Matching rules (either **permit** or **deny**) contained in the **if-match** clause (such as ACLs or IP prefix lists)

  The matching results are listed in Table 16-46.

Table 16-46 Matching results of a Route-Policy

Rule (Matching Rule Contained in if-match Clauses)	Mode (Matching Mode of a Node)	Matching Result
permit	**permit**	• Routes matching the **if-match** clauses of the node match the Route-Policy, and the matching is complete. • Routes not matching the **if-match** clauses of the node continue to match the next node of the Route-Policy.
	deny	• Routes matching the **if-match** clauses of the node are denied by the Route-Policy, and the matching is complete. • Routes not matching the **if-match** clauses of the node continue to match the next node of the Route-Policy.
deny	**permit**	• Routes matching the **if-match** clauses of the node are denied by the Route-Policy and continue to match the next node. • Routes not matching the **if-match** clauses of the node continue to match the next node of

Rule (Matching Rule Contained in if-match Clauses)	Mode (Matching Mode of a Node)	Matching Result
		the Route-Policy.
	deny	• Routes matching the **if-match** clauses of the node are denied by the Route-Policy and continue to match the next node. • Routes not matching the **if-match** clauses of the node continue to match the next node of the Route-Policy. **NOTE** If all **if-match** clauses and nodes of the Route-Policy are in the **deny** mode, all the routes to be filtered are denied by the Route-Policy.

NOTE

On the VRP, all unmatched routes are denied by the Route-Policy by default. If more than one node is defined in a Route-Policy, at least one of them must be in the **permit** mode. The reason is as follows:

- If a route fails to match any of the nodes, the route is denied by the Route-Policy.
- If all the nodes in the Route-Policy are set in the **deny** mode, all the routes to be filtered are denied by the Route-Policy.

Other Functions

In addition to the preceding functions, routing policies have an enhanced feature: BGP to IGP.

In some scenarios, when an IGP uses a routing policy to import BGP routes, route attributes, the cost for example, can be set based on private attributes such as the community in BGP routes. However, without the BGP to IGP feature, BGP routes are denied because the IGP fails to identify private attributes such as community attributes in these routes. As a result, **apply** clauses used to set route attributes do not take effect.

With the BGP to IGP feature, route attributes can be set based on private attributes, such as the community and AS_Path attributes in BGP routes. The BGP to IGP implementation process is as follows:

- When an IGP imports BGP routes through a routing policy, route attributes can be set based on private attributes such as the community attribute in BGP routes.
- If BGP routes carry private attributes such as community attributes, the system uses the private attributes to filter the BGP routes. If the BGP routes meet the matching rules, the routes match the routing policy, and **apply** clauses take effect.
- If BGP routes do not carry private attributes such as community attributes, the BGP routes mismatch the routing policy and are denied, and **apply** clauses do not take effect.

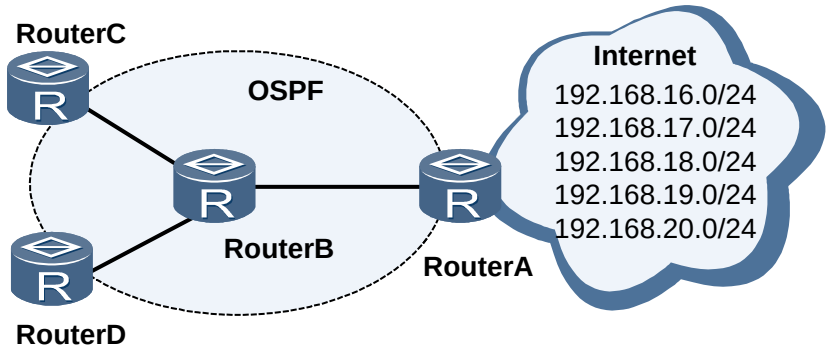
16.13.4 Applications

Specific Routes Filtering

On the OSPF-enabled network shown in Figure 16-90, Router A receives routes from the Internet and advertises some of the routes to Router B.

- Router A advertises only routes 192.168.17.0/24, 192.168.18.0/24, and 192.168.19.0/24 to Router B.
- Router C accepts only the route 192.168.18.0/24.
- Router D accepts all the routes advertised by Router B.

Figure 16-90 Networking diagram for filtering received and advertised routes



There are multiple approaches to meet the preceding requirements, and the following two approaches are used in this example:

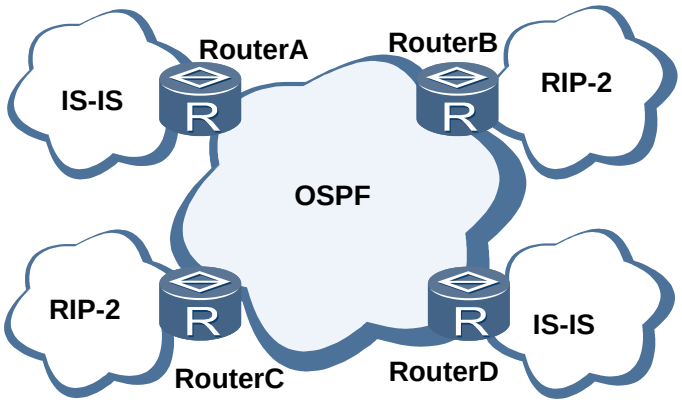
- Use IP prefix lists
  - Configure an IP prefix list for Router A and configure the IP prefix list as an export policy of Router A to be used by OSPF.
  - Configure another IP prefix list for Router C and configure the IP prefix list as an import policy of Router C to be used by OSPF.
- Use route-policies
  - Configure a Route-Policy (the matching rules can be the IP prefix list, cost, or route tag) for Router A and configure the Route-Policy as an export policy of Router A to be used by OSPF.
  - Configure another Route-Policy for Router C and configure the Route-Policy as an import policy of Router C to be used by OSPF.

  Compared with an IP prefix list, a Route-Policy allows route attributes to be modified and can be used to control routes more flexibly, but it is more complex to configure.

Transparent Transmission of Routes of Other Protocols Through an OSPF AS

On the network shown in Figure 16-91, an AS runs OSPF and functions as a transit AS for other areas. Routes from the IS-IS area connected to Router A need to be transparently transmitted through the OSPF AS to the IS-IS area connected to Router D. Routes from the RIP-2 area connected to Router B need to be transparently transmitted through the OSPF AS to the RIP-2 area connected to Router C.

Figure 16-91 Networking diagram for transparently transmitting routes of other protocols through an OSPF AS

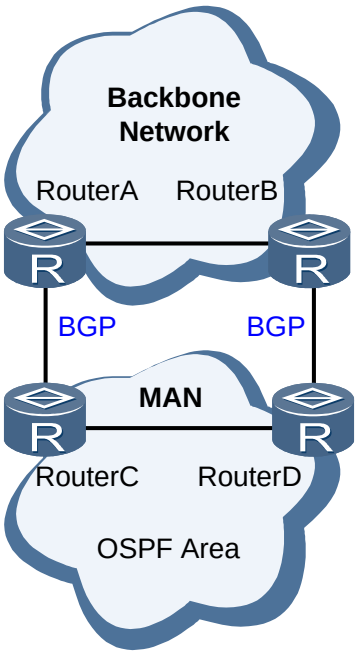


To meet the preceding requirements, configure a Route-Policy for Router A to set a tag for the imported IS-IS routes. Router D identifies the IS-IS routes from OSPF routes based on the tag.

Application of BGP to IGP

On the network shown in Figure 16-92, Router A and Router B are aggregation devices on a backbone network, and Router C and Router D are egress devices of a MAN. BGP peer relationships are established between Router A and Router C as well as between Router B and Router D. External routes are advertised to the MAN using BGP. The MAN runs OSPF to implement interworking.

Figure 16-92 Networking diagram for BGP to IGP



To enable devices on the MAN to access the backbone network, Router C and Router D need to import routes. When OSPF imports BGP routes, a routing policy can be configured to control the number of imported routes based on private attributes (such as the community) of the imported BGP routes or modify the cost of the imported routes to control the MAN egress traffic.

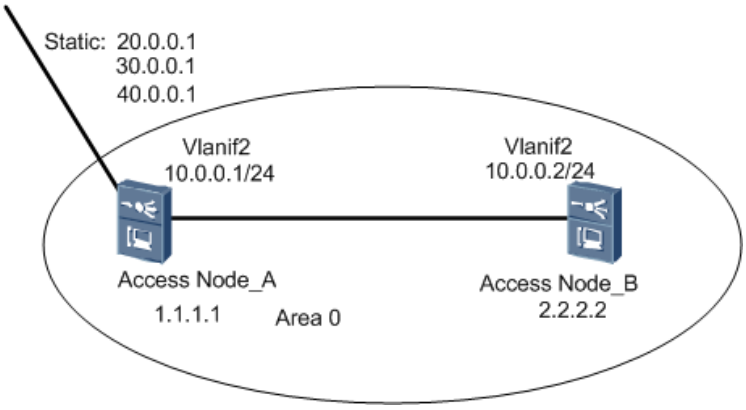
16.13.5 Configuration Example of the Routing Policy

This topic provides an example for configuring a routing policy for imported routes.

Service Requirements

- Consider two MA5600T/MA5603T/MA5608Ts with routing function enabled, namely MA5600T/MA5603T/MA5608T_A and MA5600T/MA5603T/MA5608T_B. Both of them are running the OSPF routing protocol, and within area 0.
- MA5600T/MA5603T/MA5608T_A imports static routes, and MA5600T/MA5603T/MA5608T_B is configured with the routing filtering policy.

Figure 16-93 Example network for configuring the routing policy



Procedure

Configuring MA5600T/MA5603T/MA5608T_A.

1. Configure the IP address of the Layer 3 interface on MA5600T/MA5603T/MA5608T_A.

```
huawei(config)#vlan 2 smart
huawei(config)#port vlan 2 0/19 0
huawei(config)#interface vlanif 2
huawei(config-if-vlanif2)#ip address 10.0.0.1 24
huawei(config-if-vlanif2)#quit
```

2. Enable OSPF on MA5600T/MA5603T/MA5608T_A and specify the area ID to which the interface belongs.

```
huawei(config)#ospf
huawei(config-ospf-1)#area 0
huawei(config-ospf-1-area-0.0.0.0)#network 10.0.0.0 0.0.0.255
huawei(config-ospf-1-area-0.0.0.0)#quit
huawei(config-ospf-1)#quit
```

3. Configure the OSPF router ID on MA5600T/MA5603T/MA5608T_A.

```
huawei(config)#router id 1.1.1.1
```

4. Configure three static routes.

```
huawei(config)#ip route-static 20.0.0.1 32 NULL 0
huawei(config)#ip route-static 30.0.0.1 32 NULL 0
huawei(config)#ip route-static 40.0.0.1 32 NULL 0
```

5. Import static routes into the OSPF routing table to improve its capability of obtaining routes.

```
huawei(config)#ospf
hawei(config-ospf-1)#import-route static
hawei(config-ospf-1)#quit
```

6. Save the data.

```
huawei(config)#save
```

Step 1 Configuring MA5600T/MA5603T/MA5608T_B.

1. Configure the IP address of the Layer 3 interface on MA5600T/MA5603T/MA5608T_B.

```
huawei(config)#vlan 2 smart
huawei(config)#port vlan 2 0/19 0
huawei(config)#interface vlanif 2
huawei(config-if-vlanif2)#ip address 10.0.0.2 24
huawei(config-if-vlanif2)#quit
```

2. Configure the ACL.

```
huawei(config)#acl 2000
huawei(config-acl-basic-2000)#rule deny source 30.0.0.0 255.255.255.0
huawei(config-acl-basic-2000)#rule permit source any
huawei(config-acl-basic-2000)#quit
```

3. Enable OSPF on MA5600T/MA5603T/MA5608T_B and specify the area id to which the interface belongs.

```
huawei(config)#ospf
huawei(config-ospf-1)#area 0
huawei(config-ospf-1-area-0.0.0.0)#network 10.0.0.0 0.0.0.255
huawei(config-ospf-1-area-0.0.0.0)#quit
huawei(config-ospf-1)#quit
```

4. Configure the OSPF router ID of MA5600T/MA5603T/MA5608T_B.

```
huawei(config)#router id 2.2.2.2
```

5. Filter imported routes.

```
huawei(config)#ospf
uawei(config-ospf-1)#filter-policy 2000 import
huawei(config-ospf-1)#quit
```

6. Save the data.

```
huawei(config)#save
```

----End

Result

1. MA5600T/MA5603T/MA5608T_A and MA5600T/MA5603T/MA5608T_B run OSPF successfully, and they can communicate well with each other.
2. After a filter is configured on MA5600T/MA5603T/MA5608T_B, parts of the three imported static routes are available while part of them is screened on

MA5600T/MA5603T/MA5608T_B. That is, routes from segments 20.0.0.0 and 40.0.0.0 are available, while the route from segment 30.0.0.0 is screened.

Configuration File

Configuration on MA5600T/MA5603T/MA5608T_A.

```
vlan 2 smart
port vlan 2 0/19 0
interface vlanif 2
ip address 10.0.0.1 24
quit
ospf
area 0
network 10.0.0.0 0.0.0.255
quit
quit
router id 1.1.1.1
ip route-static 20.0.0.1 32 NULL 0
ip route-static 30.0.0.1 32 NULL 0
ip route-static 40.0.0.1 32 NULL 0
ospf
import-route static
quit
save
```

Configuration on MA5600T/MA5603T/MA5608T_B.

```
vlan 2 smart
port vlan 2 0/19 0
interface vlanif 2
ip address 10.0.0.2 24
acl 2000
rule deny source 30.0.0.0 255.255.255.0
rule permit source any
quit
ospf
area 0
network 10.0.0.0 0.0.0.255
quit
quit
router id 2.2.2.2
ospf
filter-policy 2000 import
quit
save
```

16.14 ECMP

Equal and Weighted Cost Multi-Path (ECMP) is a technique in which if two or more equal cost shortest paths exist between two nodes, the traffic between the nodes is distributed among the multiple equal-cost paths.

16.14.1 Introduction to ECMP

Definition

Equal and Weighted Cost Multi-Path (ECMP) is a technique in which if two or more equal cost shortest paths exist between two nodes, the traffic between the nodes is distributed among the multiple equal-cost paths. That is, in packet transmissions, if different routes with the same destination network exist in the system, the packets can be transmitted to the destination network through multiple next hops.

Purpose

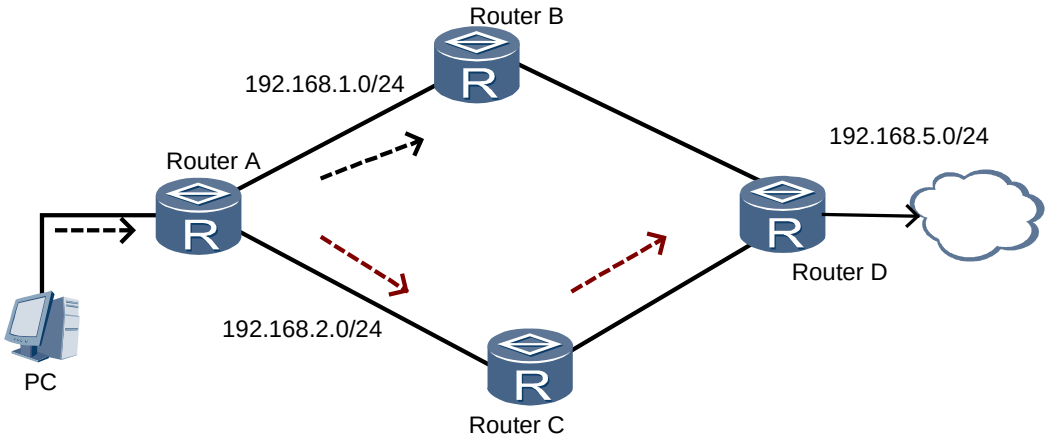
In ECMP, the traffic to the same destination network can be distributed among multiple equal-cost paths to reduce the network load, and the links in the network can back up each other. That is, when a link in the network fails, the packets on this link can be forwarded to the destination network through other links that are in the normal state.

16.14.2 ECMP Principle

In ECMP, according to different states of the network, the traffic to the same destination network can be distributed among multiple equal-cost paths to reduce the network load or to implement the link backup function.

As shown in Figure 16-94, assume that a packet is transmitted to the destination network (192.168.5.0) through Routers A-D, and two routes to the destination network exist in Router A. When receiving the packet from a user, Router A can select Router B or Router C as the next hop to forward the packet to the destination network.

Figure 16-94 ECMP diagram



17 IPv6

About This Chapter

Internet Protocol Version 6 (IPv6), also called IP Next Generation (IPng), is a second-generation protocol of the network layer protocol family.

17.1 Why IPv6 is Required

Definition

Internet Protocol Version 6 (IPv6), complies with a set of specifications defined by the Internet Engineering Task Force (IETF), IPv6 is an upgrade of Internet Protocol Version 4 (IPv4). The most significant difference between IPv6 and IPv4 is that the length of IP addresses is extended from 32 bits to 128 bits in IPv6. Featuring a simplified header format, sufficient address space, layered address structure, flexible extension header, and enhanced neighbor discovery (ND) mechanism, IPv6 is more competitive in the future market.

Purpose

As the IPv4-based Internet achieves great success, the IP technology is widely applied. With the rapid development of the Internet, however, deficiencies of IPv4 are more strongly felt, especially in the following aspects:

- The IPv4 address space is insufficient.

  An IPv4 address is identified by using 32 bits. In theory, a maximum of 4.3 billion addresses can be provided. In practice, less than 4.3 billion addresses are available due to address allocation reasons. In addition, IPv4 address resources are allocated unevenly. The USA uses almost half of the world's IP addresses, while Europe uses less IP addresses than the USA and the Asian-Pacific region uses even less. The development of mobile IP and broadband technologies requires more IP addresses. The shortage of IPv4 addresses directly restricts further development of the IP technology.

  There are several solutions to IPv4 address shortage. Classless Interdomain Routing (CIDR) and Network Address Translation (NAT) are two major solutions. CIDR and NAT, however, have their disadvantages and outstanding problems. This also drives the need for and the development of IPv6.
- The backbone device needs to maintain a large number of routing entries.

Due to allocation and planning problems in the early phase of IPv4 development, many discontinuous IPv4 addresses are allocated and as a result routes cannot be aggregated effectively. The increasingly large routing tables consume a lot of memory, leading to higher costs for equipment and lower forwarding efficiency. To tackle these issues, device manufacturers have to constantly upgrade their products in order to improve route addressing and forwarding performance.

- Address autoconfiguration and readdressing cannot be performed easily.

  An IPv4 address occupies only 32 bits and IP addresses are allocated unevenly. Consequently, IP addresses need to be reallocated during network expansion or network replanning. Address autoconfiguration and readdressing are required to simplify maintenance. With IPv4, however, such tasks cannot be performed easily.

- Security cannot be well guaranteed.

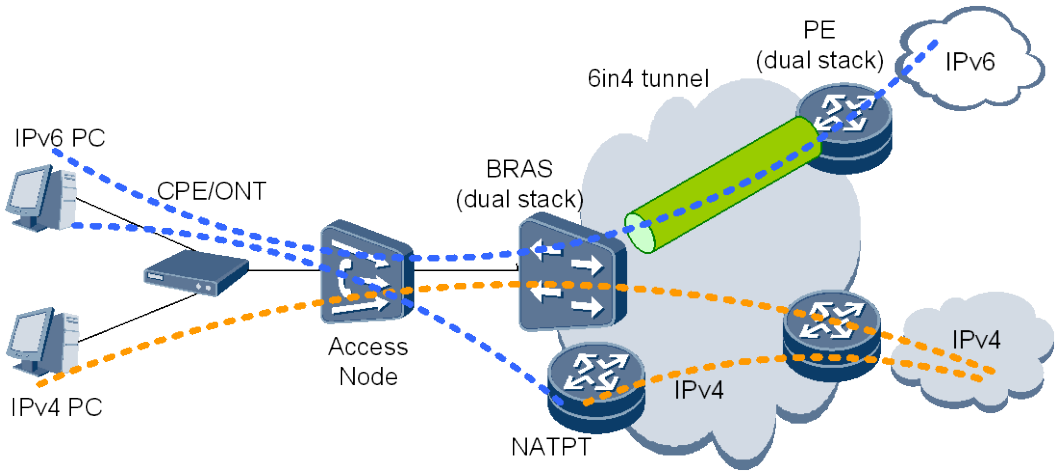
  With the development of the Internet, security problems become more serious. During the design of IPv4, security was not fully considered. Hence, the original framework cannot implement end-to-end security. IPv6 provides end-to-end security.

IPv6 solves the IP address shortage issue from the very outset. Moreover, IPv6 is easy to deploy, compatible with various applications, allows IPv4 networks to smoothly transit to IPv6 networks, and can coexist and interwork with IPv4. With so many advantages over IPv4, IPv6 is in rapid development.

17.2 IPv6 network deployment

Internet protocol version 6 (IPv6) is developed to tackle global IPv4 address shortage. At the early stage of IPv4-to-IPv6 transition, IPv4 networks have been in large-scale deployment whereas deployment of IPv6 networks is comparatively scarce, and both types of networks will coexist. Figure 17-1 shows the IPv6 network deployment at the early stage of IPv4-to-IPv6 transition.

Figure 17-1 Early-stage IPv6 network deployment



IPv4-based broadband remote access servers (BRASs) on the live network are progressively upgraded to support IPv4/IPv6 dual-stack. The dual-stack BRAS provides a 6in4 tunnel or a dedicated link to transmit IPv6 traffic to an IPv6 network.

The dual-stack BRAS provides the network address translation-protocol translation (NAT-PT) function to allow IPv6 users to access an IPv4 network.

The access network equipment is able to detect IPv6 packets, allocate IPv6 addresses (using DHCPv6), and support ACLv6.

17.3 Principles

Basic functions of IPv6 include IPv6 neighbor discovery and IPv6 path maximum transmission unit (PMTU) discovery. Neighbor discovery and PMTU discovery are implemented through Internet Control Message Protocol for IPv6 (ICMPv6) messages.

17.3.1 IPv6 Highlights

- A 128-bit address structure, providing sufficient address space

  A major advantage of IPv6 is the almost infinite IP address space. IPv6 increases the size of an IP address from 32 bits to 128 bits, which is four times of that of IPv4. A 128-bit address structure is able to provide about $4{,}300{,}000{,}000^4$ addresses, meeting almost any address assignment requirements that can be predicted.

- Layered address structure

  The layered address structure realizes rapid route lookup, reduces the size of IPv6 routing tables with the aid of route aggregation, and thereby improves the forwarding efficiency of routers.

- Address autoconfiguration

  IPv6 enables hosts to discover networks and obtain IPv6 addresses using address autoconfiguration, which greatly improves the network manageability. Using address autoconfiguration, user devices (such as mobile phones and wireless devices) support plug-and-play, without requiring manual configuration or using a private server (such as a DHCP server). IPv6 supports stateful address autoconfiguration and stateless address autoconfiguration.

  - In stateful address autoconfiguration, the host obtains the address and other configuration information from the server.
  - In stateless address autoconfiguration, the host automatically configures address information that contains the prefix and interface ID of the host as reported by the local router. If there is no router on the link, the host can automatically configure only a link-local address for interoperating with the local node.

- Source/Destination address selection

  To specify or plan source/destination addresses of the packets sent by the system, the network administrator can define a set of address selection rules. These rules form an address selection policy table. The policy table is similar to a routing table and employs the longest matching rule for prefix lookup. The address selection result is determined together by the source address and destination address.

  A source address is selected according to the following rules. Among the rules below, the rule with a smaller number has a higher priority. A candidate address is preferred if it:

  a. Is the same as the destination address.
  b. Has an appropriate effective scope.
  c. Is not a deprecated address.
  d. Is a home address.

e. Is the address of an outgoing interface.

f. Has the same *label* value as the destination address.

g. Has the longest matching prefix.

NOTE

The candidate address can be a unicast address that is configured on the specified outgoing interface. If a source address that has the same *label* value as and is in the same address range with the destination address is not found on the outgoing interface, such a source address can be selected on another interface.

A destination address is selected according to the following rules. Among the rules below, the rule with a smaller number has a higher priority. A candidate address is preferred if it:

a. Is not an unusable address.

b. Has an appropriate effective scope.

c. Is not a deprecated address.

d. Is a home address.

e. Has the same *label* value as the source address.

f. Has a larger *precedence* value.

g. Prefer native transport (6over4 or 6to4 tunnel is not required).

h. Has a smaller effective scope.

i. Has the longest matching prefix.

j. Otherwise, leave the order unchanged.

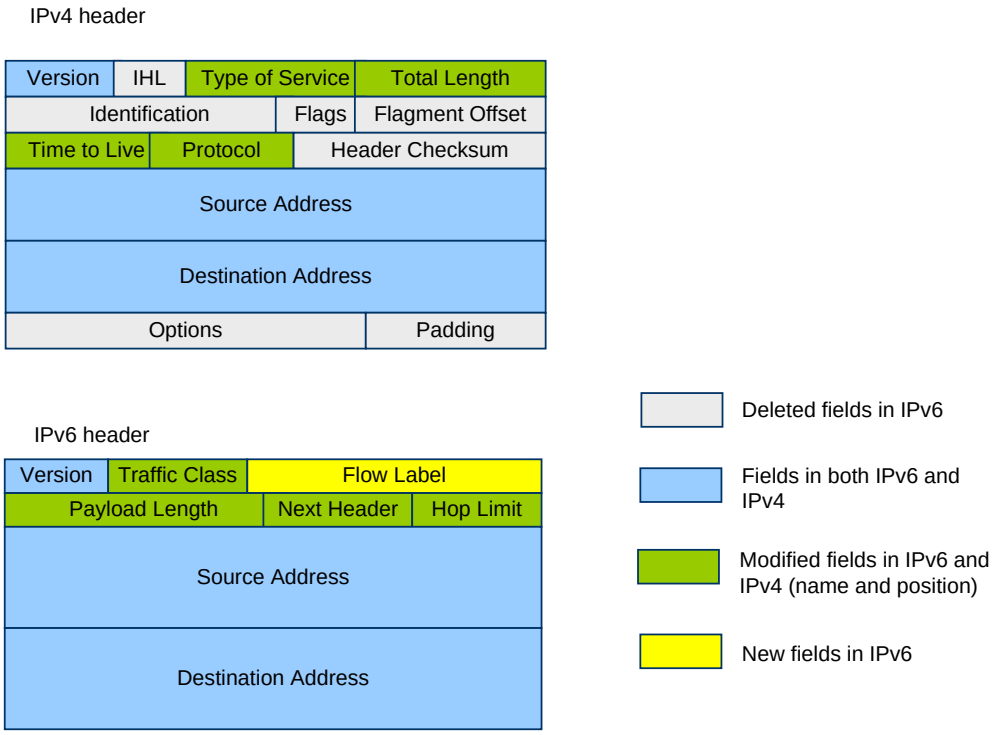
- Support for QoS

  New fields are added in the IPv6 header to define how to label and process flows. Flows are identified by the Flow Label field in the header. The Flow Label field allows routers to identify the packets of a certain flow and provide special processing for these packets.

- Flexible and simple extension headers

  Figure 17-2 shows the comparison between IPv6 and IPv4 header formats. Compared with an IPv4 header, an IPv6 header deletes the IHL, Identification, Flags, Fragment Offset, Header Checksum, Options, and Padding fields, and adds the Flow Label field, which helps improve the header processing efficiency. In addition, to better support different options, IPv6 introduces multiple extension headers. With these extension headers, it is unnecessary to modify the existing packet structure when new options are added, which greatly improves the flexibility of IPv6.

Figure 17-2 Comparison between IPv6 and IPv4 header formats



17.3.2 IPv6 Addresses

Format of an IPv6 Address

A 128-bit IPv6 address has two formats:

- X:X:X:X:X:X:X:X
  - In this format, the 128 bits of an IPv6 address are divided into 8 groups. The 16 bits of each group are represented by 4 hexadecimal characters (0 to 9, and A to F). Groups are separated by colons. Every "X" represents the numerical value of a group of hexadecimal characters. The following is an example:

    2031:0000:130F:0000:0000:09C0:876A:130B

    For convenience, the zeros at the beginning of each group can be omitted. The preceding example, thus, can be written as 2031:0:130F:0:0:9C0:876A:130B.
  - Furthermore, the two or more consecutive zeros in the address can be replaced by "::", which reduces the written length of an IPv6 address. The preceding example can be further compressed as 2031:0:130F::9C0:876A:130B.

    An IPv6 address contains only one "::". Otherwise, a computer cannot determine the count of zeros when restoring the original 128-bit address from the compressed address.
- X:X:X:X:X:X:d.d.d.d

  Addresses in this format fall into the following two types:
  - IPv4-compatible IPv6 address: The format of an IPv4-compatible IPv6 address is 0:0:0:0:0:0:IPv4-address. The high-order 96 bits are all 0s, and the low-order 32

bits are an IPv4 address. This IPv4 address must be reachable in an IPv4 network, and cannot be a multicast, broadcast, loopback, or unspecified address (0.0.0.0).

- IPv4-mapped IPv6 address: The format of an IPv4-mapped IPv6 address is 0:0:0:0:0:FFFF:IPv4-address. This type of IPv6 address is used to represent the addresses of IPv4 nodes.

An IPv4-compatible IPv6 address is used for the configuration of IPv6 over IPv4 tunnels.

"X:X:X:X:X:X" represent the high-order 6 groups of numbers, and each "X" stands for 16 bits expressed in hexadecimal notation. "d.d.d.d" represent the low-order 4 groups of numbers, and each "d" stands for 8 bits expressed in decimal notation. "d.d.d.d" is a standard IPv4 address.

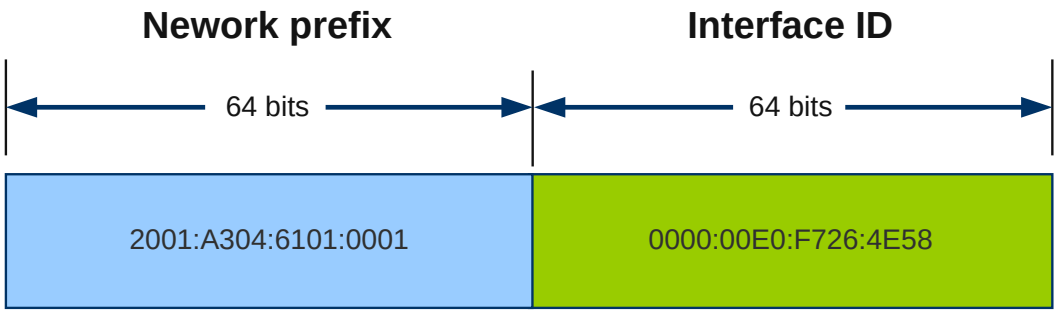
Structure of an IPv6 Address

An IPv6 address can be divided into two parts:

- Network prefix: n bits, equaling the network ID in an IPv4 address.
- Interface identifier (ID): 128-n bits, equaling the host ID in an IPv4 address.

Figure 17-3 illustrates the structure of the address 2001:A304:6101:1::E0:F726:4E58 /64.

Figure 17-3 Structure of the address 2001:A304:6101:1::E0:F726:4E58 /64



IPv6 Address Classification

IPv6 has the following types of addresses:

- Unicast address: uniquely identifies an interface and is similar to an IPv4 unicast address. The packets sent to a unicast address are transmitted to the unique interface identified by this address.

  Unicast addresses can be classified into the following types, as shown in Table 17-1.

Table 17-1 Types of IPv6 unicast addresses

Address Type	Binary Prefix	IPv6 Prefix Identifier
Link-local unicast address	1111111010	FE80::/10
Loopback address	00...1 (128 bits)	::1/128
Unspecified address	00...0 (128 bits)	::/128

Address Type	Binary Prefix	IPv6 Prefix Identifier
Global unicast address	Others	-

The meanings of each type of address are described as follows:

- Link-local IPv6 unicast address: It is used in the Neighbor Discovery (ND) protocol, and in the communication between nodes on the local link during the stateless address autoconfiguration. The packets carrying the link-local IPv6 unicast address as the source or destination address are forwarded only on the local link. A link-local IPv6 unicast address can be automatically configured on any interface by using the link-local prefix FE80::/10(1111 1110 10) and the interface identifier in the IEEE EUI-64 format (an EUI-64 identifier can be derived from an EUI-48 identifier).
- Loopback address: It is 0:0:0:0:0:0:0:1 or ::1 and not assigned to any interface. Similar to the case of IPv4 loopback address 127.0.0.1, IPv6 packets carrying the IPv6 loopback address as the destination address are looped back to the sending node.
- Unspecified address (::): It cannot be assigned to any node or function as a destination address. The unspecified address can be used in the Source Address field of the IPv6 packet sent by an initializing host that has not obtained its own address. During duplicate address detection (DAD), the Source Address field of a Neighbor Solicitation (NS) message is an unspecified address.
- Global unicast address: It is equivalent to an IPv4 public network address. Global unicast addresses are used on the links that can be aggregated, and are provided to Internet service providers (ISPs). The structure of this type of address allows for route prefix aggregation to relieve the global routing entry resources limitation. A global unicast address consists of a 48-bit route prefix that is managed by the carrier, a 16-bit subnet ID that is managed by the local node, and a 64-bit interface ID. Unless otherwise specified, global unicast addresses include site-local unicast addresses.

- Anycast address: identifies a group of interfaces, which generally belong to different nodes. The packets carrying an anycast destination address are transmitted to the interface that is nearest to the source node in the interface group identified by the anycast address. The nearest interface refers to the interface with the smallest distance metric measured by the routing protocol.

  Application scenario: When a mobile host needs to communicate with the mobile agent on the home subnet, the mobile host uses the anycast address of the routing device in the subnet.

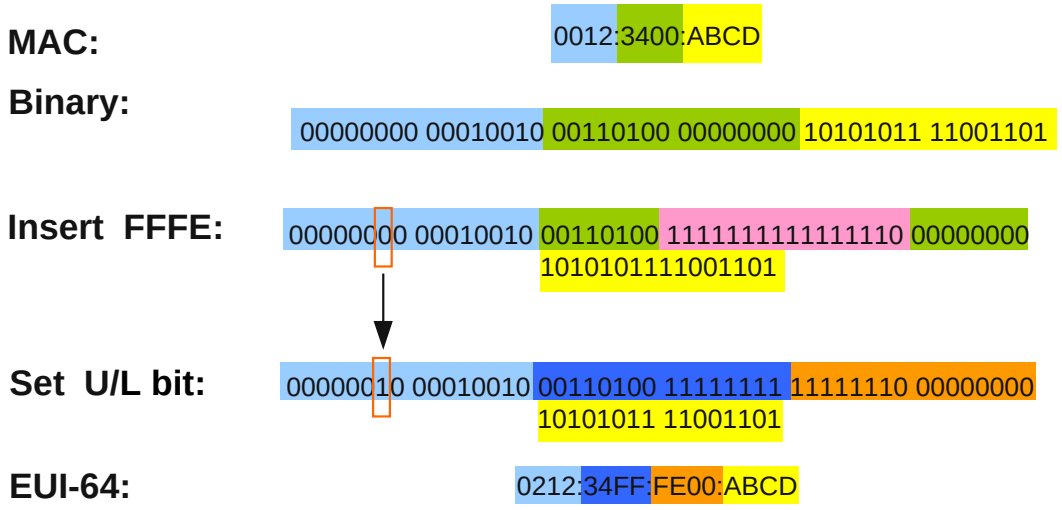
  Specifications of anycast addresses: Anycast addresses do not have independent address space. They can use the format of any unicast address. Thus, a syntax is required to differentiate an anycast address from a unicast address.
- Multicast address: identifies a group of interfaces that belong to different nodes and is similar to an IPv4 multicast address. The packets carrying a multicast destination address are transmitted to all the interfaces identified by this multicast address.

  IPv6 addresses do not include broadcast addresses. In IPv6, functions of broadcast addresses are provided by multicast addresses.

Interface ID in the IEEE EUI-64 Format

The 64-bit interface ID in an IPv6 address identifies a unique interface on a link. This address is derived from the link-layer address (such as a MAC address) of the interface. The 64-bit IPv6 interface ID is transformed from a 48-bit MAC address by inserting a hexadecimal number FFFE (1111 1111 1111 1110) into the MAC address and then setting the U/L bit (the leftmost seventh bit) to 1. Figure 17-4 shows transformation from a MAC address to an EUI-64 address.

Figure 17-4 Transformation from a MAC address to an EUI-64 address

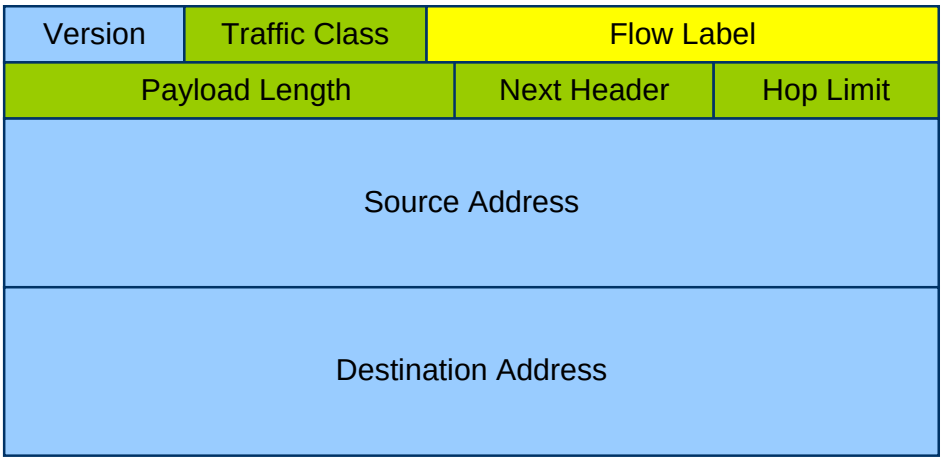


17.3.3 IPv6 Packet Format

Format of an IPv6 Header

Figure 17-5 shows the format of an IPv6 header.

Figure 17-5 Format of an IPv6 header



- Version:

  4 bits. The value of this field is 6, indicating an IPv6 packet.

- Traffic Class:

  8 bits. This field is similar to the TOS field in an IPv4 header.

- Flow Label:

  20 bits. This field is new in IPv6. Flow labels are used to label particular flows so as to differentiate packets at the network layer. Routers along a forwarding path differentiate and process packets according to the flow labels. Because the Flow Label field locates in the IPv6 header, forwarding routers and destination nodes do not need to differentiate packets according to the packet content.

- Payload Length:

  16 bits. This field indicates the length of the IPv6 payload in octets. The payload is the rest of the packet following the IPv6 header. (Note that any extension headers present are considered part of the payload.)

- Next Header:

  8 bits. This field identifies the type of header immediately following the current IPv6 header (which may be a header or an extension header). This field uses the same values as the IPv4 Protocol field. The Next Header field in the IPv6 header and the Next Header fields in the IPv6 extension headers form a chain. The chain mechanism helps improve the efficiency of extension header processing because the router processes only the option headers needed.

- Hop Limit:

  8 bits. This field is similar to the IPv4 TTL field. This field decreases by 1 by each node that forwards the packet. The packet is dropped if this field decreases to 0.

- Source Address:

  128 bits. This field indicates the source address of the packet.

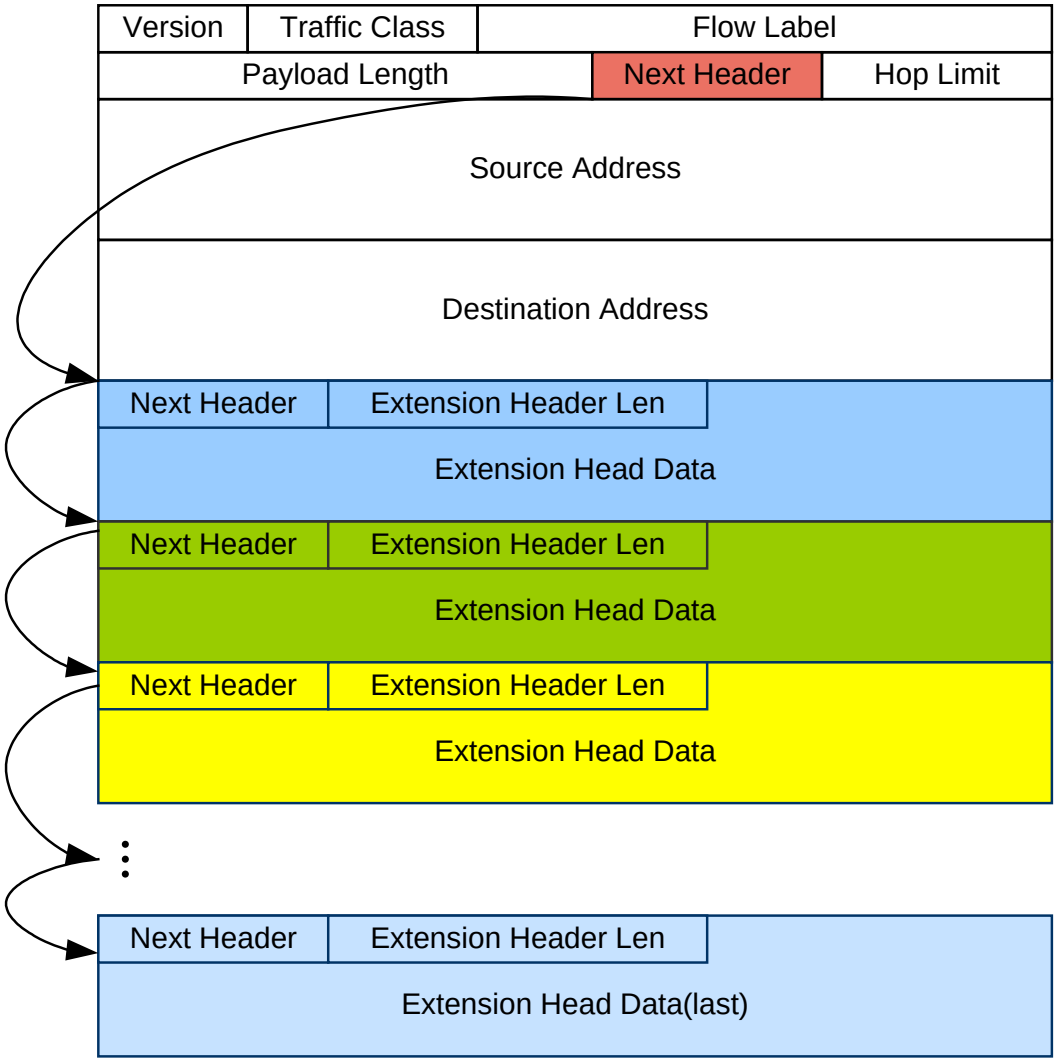
- Destination Address:

  128 bits. This field indicates the destination address of the packet.

Format of IPv6 Extension Headers

Figure 17-6 shows the format of IPv6 extension headers.

Figure 17-6 Format of IPv6 extension headers



IPv6 option fields are supported through a chain of extension headers. An IPv6 packet can carry zero, one, or multiple extension headers.

IPv6 extension headers appear in the following order:

- Hop-by-Hop Options Header

  The value of this header is 0, which is defined in the IPv6 header. It is used for routing alarms (RSVP and MLDv1) and jumbo frames. This header is processed by every node along the packet forwarding path.
- Destination Options Header

  The value of this header is 60. This header may occur before the following two headers:
  - Routing Header

    In such a case, the Destination Options header is processed by the destination node and the node specified in the Routing header.
  - Upper-layer Header (located behind any ESP option)

    In such a case, the Destination Options header is processed only by the destination node. The Destination Options header is used in the mobile IPv6 scenario.

- Routing Header

  The value of this header is 43. This header is used for source routing options and mobile IPv6.

- Fragment Header

  The value of this header is 44. This header is used for packet fragmentation when the packet sent by the source node is larger than the path maximum transmission unit (PMTU). PMTU is the MTU specified for the path from the source node to destination node.

- Authentication Header

  The value of this header is 51. This header is used for authentication and integrity checking of a packet. The definition of this header in IPv6 is the same as that in IPv4.

- ESP Header

  The value of this header is 50. This header is used for authentication, integrity checking, and encryption of a packet. The definition of this header in IPv6 is the same as that in IPv4.

- Upper-layer Header

  This header is an upper-layer protocol (such as TCP, UDP, or ICMP) header.

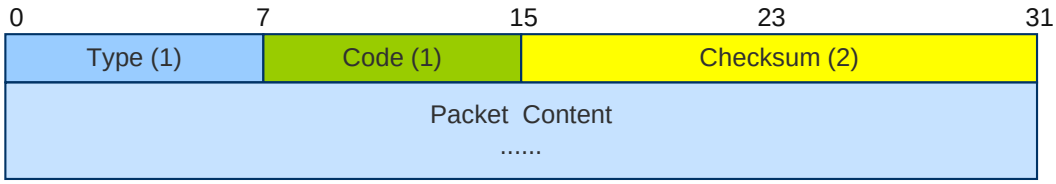
The Destination Options header occurs at most twice (once before the Routing header and once before the upper-layer header). The other extension headers occur at most once.

IPv6 nodes must be able to process the extension headers, regardless of the order and the number of occurrences they appear in the same packet, except for the Hop-by-Hop Options header that is restricted to appear immediately after an IPv6 header only. This requirement ensures interoperability between IPv6 nodes.

17.3.4 ICMPv6

As a basic protocol of IPv6, Internet Control Message Protocol for IPv6 (ICMPv6) is used by a device to generate two types of messages: error messages and informational messages. Using these messages, IPv6 nodes report errors and information generated during packet processing. Figure 17-7 shows the format of an ICMPv6 message.

Figure 17-7 Format of an ICMPv6 message



The meaning of each field in an ICMPv6 message is as follows:

- Type field: indicates the message type. The values from 0 to 127 indicate the error message type, and values from 128 to 255 indicate the informational message type.
- Code field: indicates the specific message type.
- Checksum field: indicates the checksum of an ICMPv6 message.

Classification of ICMPv6 Error Messages

- Destination Unreachable message

  When an IPv6 node is forwarding IPv6 packets and finds that the destination address of the packets is unreachable, it sends an ICMPv6 Destination Unreachable message to the source node of the packets. Specific causes for the error message are carried in the message. Destination Unreachable messages are classified into types that include the following:

  - No route to destination
  - Address unreachable
  - Port unreachable

- Packet Too Big message

  When an IPv6 node is forwarding IPv6 packets and finds that the size of the packets exceeds the path maximum transmission unit (PMTU) of the outbound interface, it sends an ICMPv6 Packet Too Big message to the source node of the packets. The PMTU of the outbound interface is carried in the message. PMTU discovery is implemented based on Datagram Too Big messages.

- Time Exceeded message

  During the reception and transmission of IPv6 packets, when a device receives a packet with the hop limit of 0 or when the device reduces the hop limit to 0, it sends an ICMPv6 Time Exceeded message to the source node of the packets. When reassembling fragmented packets, an ICMPv6 Time Exceeded message is also generated if the reassembly time exceeds the specified duration.

- Parameter Problem message

  When a destination node receives an IPv6 packet, it checks the validity of the packet. If the destination node detects any of the following errors, it sends an ICMPv6 Parameter Problem message to the source node of the packet:

  - A field in the IPv6 header or extension header is incorrect.
  - The Next Header field in the IPv6 header or extension header cannot be identified.
  - Unknown options exist in the extension header.

Classification of ICMPv6 Informational Messages

ICMPv6 informational messages are classified into Echo Request messages and Echo Reply messages. ICMPv6 informational messages can be used for network fault diagnosis, PMTU discovery, and neighbor discovery. During the interoperation check between two nodes, the node that receives an Echo Request message sends an Echo Reply message to the source node. In this manner, subsequent packets are received and transmitted between the two nodes.

17.3.5 PMTU

Problems Related to MTU

A path from a source address to a destination address may traverse interfaces that have different maximum transmission unit (MTU) values. The smallest MTU on this path is called the path MTU (PMTU).

- During transmission of IPv6 packets, the packets cannot be fragmented on the intermediate nodes. Therefore, it often happens that the packet length exceeds the PMTU. In such a case, the source node needs to retransmit the IPv6 packets, which reduces the transmission efficiency.

- If the source node uses the minimum link MTU (1280 bytes) as the maximum fragment length, the fragments sent by a node will always be smaller than the PMTU, because in most cases the PMTU is greater than the minimum link MTU of the path. This results in network resources wasting.

The PMTU discovery protocol is introduced to solve this problem.

Principle of PMTU Discovery

PMTU discovery is the process of discovering the optimal link MTU on the path from the source to the destination. PMTU discovery describes a method of dynamically discovering the PMTU for a path. When an IPv6 node sends a large amount of data to another node, the data is transmitted by means of a sequence of IPv6 fragments. When these fragments are of the maximum length allowed in successful transmission between the source node and destination node, the fragment length is considered optimal and called PMTU.

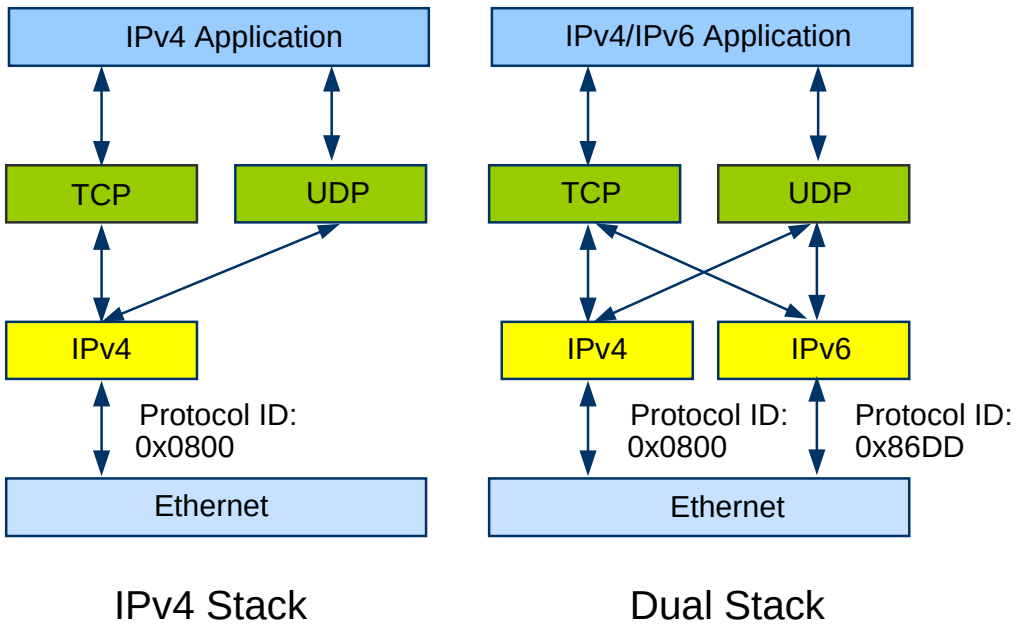
A source node assumes that a PMTU of a path is the known IPv6 MTU of the first hop on the path. If the packet sent from this path is too large to be forwarded along the path, the intermediate node discards this packet and returns an ICMPv6 Packet Too Big message to the source node. The source node then sets the PMTU of the path according to the IPv6 MTU specified in the message.

When the PMTU learned by the node is smaller than or equal to the actual PMTU, the PMTU discovery process ends. Before the PMTU discovery process end, ICMPv6 Packet Too Big messages may be repeatedly sent and received because smaller IPv6 MTUs may be found further down the path.

17.3.6 Dual Protocol Stacks

For an IPv6 node, the most effective way of being compatible with IPv4 is to retain a complete IPv4 protocol stack on the node. Such a node is called a dual-stack node. Figure 17-8 shows the structures of a single protocol stack and dual protocol stacks.

Figure 17-8 Structures of a single protocol stack and dual protocol stacks in an Ethernet



Dual protocol stacks have the following advantages:

- Multiple link-layer protocols support dual protocol stacks.

  Multiple link-layer protocols, such as Ethernet, support dual protocol stacks. In Figure 17-8, the link-layer protocol is Ethernet. In an Ethernet frame, if the Protocol ID field is 0x0800, it indicates that the network layer receives IPv4 packets; if the field is 0x86DD, it indicates that the network layer receives IPv6 packets.

- Multiple applications support dual protocol stacks.

  The upper layer applications, such as the DNS, can use TCP or UDP as the transmission layer protocol, and prefers the IPv6 protocol stack rather than the IPv4 protocol stack as the network-layer protocol.

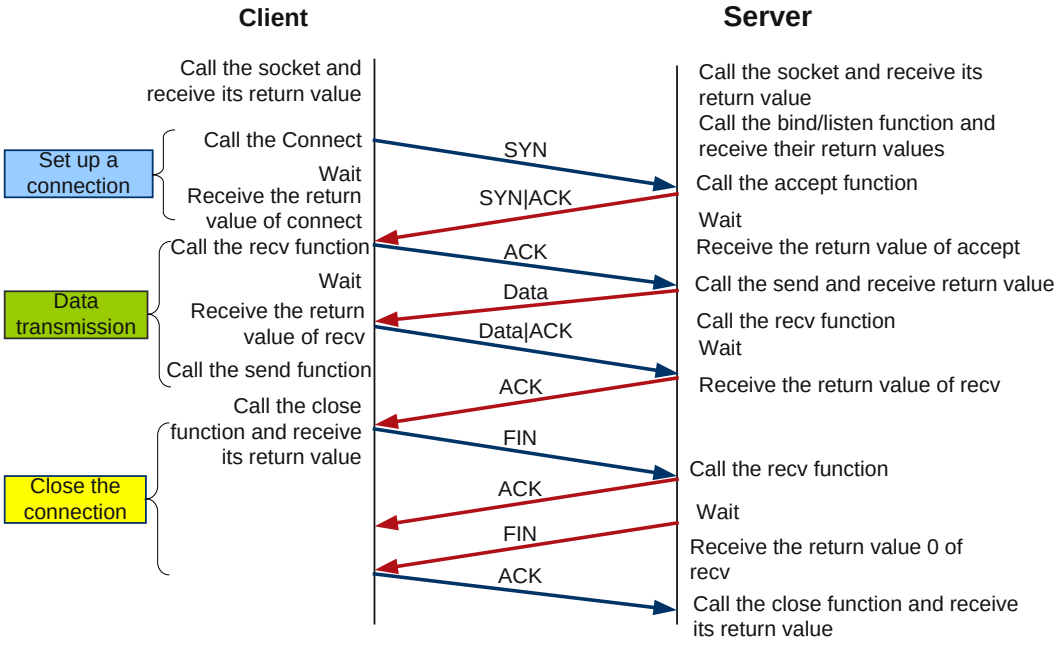
17.3.7 TCP6

Transmission Control Protocol Version 6 (TCP6) provides a mechanism to establish virtual circuits between processes of two endpoints. A TCP6 virtual circuit is similar to the full-duplex circuit that transmits data between systems. TCP6 is called a reliable protocol because it provides reliable data transmission between processes. TCP6 also provides a mechanism to optimize the transmission performance according to the network status. When all the data can be received and acknowledged, the transmission rate increases gradually.

TCP6 is generally used in interactive applications, such as the web application. TCP6 establishes virtual circuits by using the three-way handshake mechanism, and all the virtual circuits are deleted using the four-way handshake mechanism. TCP6 connections provide a variety of checksums and reliability-ensuring functions, but increase the cost. As a result, TCP6 has lower efficiency than User Datagram Protocol Version 6 (UDP6).

Figure 17-9 shows the establishment and removal of a TCP6 connection.

Figure 17-9 Establishment and removal of a TCP6 connection



17.3.8 UDP6

User Datagram Protocol Version 6 (UDP6) is a computer communication protocol used to exchange packets on a network. UDP6 has the following characteristics:

- UDP uses only source and destination information and runs mainly in a simple request/response structure.
- UDP is unreliable. No control mechanism is provided to ascertain whether UDP6 datagrams have reached their destinations.
- UDP is connectionless. No virtual circuits are required for data transmission between hosts.

The connectionless characteristics of UDP6 enables UDP6 to send data to broadcast addresses. This is different from TCP6, which requires specific source and destination addresses.

17.3.9 RawIP6

RawIP6 is an implementation in which only a limited number of fields in the IPv6 header are filled with values, and RawIP6 allows applications to provide their own IPv6 headers.

RawIP6 is similar to UDP6 in the following aspects:

- RawIP6 is unreliable. No control mechanism is provided to ascertain whether RawIP6 datagrams have reached their destinations.
- RawIP6 is connectionless. No virtual circuits are required for data transmission between hosts.

Unlike UDP6, RawIP6 allows applications to directly operate the IP layer through sockets. Therefore, RawIP6 is convenient for the applications that need to interact directly with the lower layer.

17.3.10 Neighbor Discovery

The Neighbor Discovery (ND) protocol defines a set of messages and processes for determining the relationship between neighboring nodes. The IPv6 ND protocol supports the Address Resolution Protocol (ARP) messages, and the Router Discovery and Redirect messages of the Internet Control Message Protocol (ICMP) of IPv4, and also provides other functions.

After an IPv6 address is configured on a node, the node first verifies that the address is available and does not conflict with other addresses. If the node is a host, the router needs to notify the host of a better first-hop address to reach a particular destination. If the node is a router, this node needs to advertise its address, address prefix, and other configuration information, which provides guidance for the host to configure parameters. During IPv6 packet forwarding, a node needs to verify the link-layer address and reachability of its neighboring node. IPv6 ND defines five types of ICMPv6 packets:

- Router Solicitation (RS): Sent by an enabled host to a router. The router then responds with a Router Advertisement (RA).
- RA: Advertised by a router periodically, which contains information such as the prefix and some flags.
- Neighbor Solicitation (NS): Sent by an IPv6 node to determine the link-layer address of a neighbor, to check whether a neighbor is reachable, or for duplicate address detection (DAD).

- Neighbor Advertisement (NA): A response to an NS. An IPv6 node may also send unsolicited NAs to announce a link-layer address change.
- Redirect: Used by a router to inform hosts of a better first hop for a specific destination when the router finds that the incoming interface and outgoing interface of the packets are the same.

The ND protocol for IPv6 has the following functions.

DAD

In an IPv6 network, a link-local address (LLA) is used for the communication between neighboring nodes on the same link, for example, the communication between hosts on a single link where no routers exist. If an LLA is available, it is automatically used for neighbor discovery. Duplicate address detection (DAD) is a detection mechanism used for determining whether an LLA address is available. The process is as follows:

1. When an IPv6 address is configured on a node, the node sends an NS to its neighboring node to check whether the IP address is already used.
2. When receiving the NS, the neighboring node checks whether it has the same IPv6 address. If yes, the neighboring node responds with an NA carrying the IPv6 address information.
3. The node, upon receiving the NA, considers that the IPv6 address configured is already used by its neighboring node. If the node does not receive any response packet from its neighboring node, the IPv6 address configured is available.

DAD Proxy

A DAD proxy is used to prevent LLA conflicts on a Layer 3 interface. The DAD proxy feature resolves the issue of LLA conflicts between isolated ports on the same Layer 3 interface.

- When a user performs DAD, the MA5600T/MA5603T/MA5608T checks whether any LLA conflict occurs in the Layer 3 interface VLAN based on the information about the packet sent by the user.
  - If no conflict occurs, the MA5600T/MA5603T/MA5608T forwards the packet.
  - If a conflict occurs, the MA5600T/MA5603T/MA5608T responds to this conflict and drops the packet.
- The MA5600T/MA5603T/MA5608T obtains the LLA lease time from Router Advertisement (RA) and DHCPv6 packets, and the buffered LLA entries age when the lease time expires.

Neighbor Discovery

ND for IPv6, similar to ARP in IPv4, is used to parse the addresses of neighbors and detect whether neighbors are reachable using NSs and NAs.

To obtain the link-layer address of another node in the same local link, a node (source node) sends an NS with its ICMPv6 type as 135. This packet is similar to an ARP request packet in IPv6; but unlike the ARP request packet using a broadcast address, the NS uses a multicast address. A node with the last 24 bits of its address the same as the multicast address will receive this NS, which reduces the possibility of broadcast storm. The node receiving the NS (destination node) fills in its link-layer address in the response packet.

An NS can also be used to check whether a neighboring node is reachable when the link-layer address of the neighboring node is known. An NA is the response packet of an NS. The destination node, upon receiving an NS, responds with an NA with its ICMPv6 type as 136 over the local link. The source node then is able to communicate with the destination node after receiving the NA. A node may also send unsolicited NAs to announce a link-layer address change on the local link.

Router Discovery

Router Discovery is used to locate neighboring routers as well as learn prefixes and configuration parameters related to stateless address autoconfiguration (SLAAC). Router Discovery in IPv6 is implemented using the following two mechanisms:

- Router Solicitation (RS) message

  When unicast addresses are not configured on a host (for example, the system just starts up), the host sends an RS. An RS facilitates the host autoconfiguration, without having to wait for the RA sent by IPv6 routers. An RS packet is an ICMPv6 packet of type 133.

- Router Advertisement (RA) message

  Each RA-enabled IPv6 router periodically sends RAs. After receiving an RS from an IPv6 node on the local link, an IPv6 router also responds with an RA. An IPv6 router sends an RA to the multicast addresses (FF02::1) of all nodes or to the IPv6 unicast address of the node sending an RS. An RA is an ICMPv6 of type 134, including the following content:

  - Whether to use address autoconfiguration
  - Supported autoconfiguration type (stateless or stateful)
  - One or more local link prefixes (nodes on the local link can use these prefixes to perform address autoconfiguration)
  - Lifetime of local link prefixes advertised
  - Whether the router sending an RA can serve as a default router. If yes, the information also contains the lifetime of the default router in units of seconds.
  - Other host-related information, such as the hop limit and the MTU used for the host to send packets

  An IPv6 node on the local link receives an RA and updates information such as the default router, prefix list, and other information from this RA.

Address Autoconfiguration

RAs and per-prefix flags enable routers to inform hosts how to perform address autoconfiguration. For example, routers can specify whether hosts use stateful (DHCPv6) or stateless (SLAAC) address configuration.

When using the SLAAC protocol, a host uses the prefix information and local interface ID obtained from the RA received to automatically generate an IPv6 address. Also, the host can set the default router according to the default router information in the RA.

Redirect

Redirect messages are sent by routers to inform a host of a better first-hop IPv6 address for a specific destination. Like IPv4, IPv6 Redirect messages are sent only for redirecting packets to a better router. Nodes receiving these Redirect messages will send packets to this better router. Routers send Redirect messages only for unicast flows, and Redirect messages are only sent to and processed by those nodes (hosts) triggering redirect.

Default Router Priority and Routing Information

An RA defines two fields: default router priority and routing information, which helps hosts select a better forwarding router for packets.

When the link where a host resides has multiple routers, the host needs to select the forwarding router according to the packet's destination address. In such cases, routers advertise default router priority and specific routing information to hosts, improving hosts' capability of selecting better forwarding routers according to different destination addresses.

A host, upon receiving an RA that contains the routing information, will update its routing list. Before sending packets to other devices, the host selects a better route according to its routing list.

A host updates its default router list after receiving an RA that contains default router priority information. Before sending packets to other devices, the host queries its default router list and selects the highest-priority router for sending packets if no route is available. If the highest-priority router is faulty, the host will choose the second-highest-priority router and so on.

ND Proxy

User-side packets are sent to only the upper-layer device and they cannot be forwarded between user sides. After multicast domains are isolated, user networks are more secured. However, user terminals require interconnection, that is, different VLANs or some users isolated in the same VLAN need to communicate with each other.

Both IPv4 and IPv6 networks have such a requirement. It is similar to ARP proxy for IPv4 networks. ND proxy is an expansion of ARP proxy for IPv6 networks, which is used to resolve the IPv6 network interconnection issue. To configure it,

- Enable the ND proxy on a super VLAN interface, and then the super VLAN interface acts as the proxy of all the sub VLANs in the super VLAN.
- To enable the ND proxy for the NEs in a sub VLAN, the ND proxy of the sub VLAN must be enabled.

17.4 Configuring Basic IPv6 Information

This topic describes the IPv6 features supported by the MA5600T/MA5603T/MA5608T. The basic IPv6 configuration includes configuration of the IPv6 address, IPv6 neighbor discovery, path maximum transmission unit (PTMU), and transmission control protocol 6 (TCP6).

Context

Internet Protocol Version 6 (IPv6), as a set of specifications defined by the Internet Engineering Task Force (IETF), IPv6 is the upgraded version of Internet Protocol Version 4 (IPv4). The most obvious difference between IPv6 and IPv4 is that IP addresses are lengthened from 32 bits to 128 bits. IPv6 radically solves the problem of IP address shortage. Moreover, IPv6 has the following advantages: It is easy to deploy, compatible with various applications, easy for IPv4 networks to transit to IPv6 networks, and coexists and interworks with IPv4 networks.

The following table lists the IPv6 features supported by the MA5600T/MA5603T/MA5608T.

Table 17-2 IPv6 features supported by the MA5600T/MA5603T/MA5608T

Feature	Sub-feature	Configuration Process or Command
IPv6 address management and assignment	Static configuration of IPv6 global unicast addresses and IPv6 link-local addresses	17.4.1 Configuring an IPv6 Address for an Interface
	Automatic configuration of IPv6 link-local addresses	ipv6 address auto link-local
	DHCPv6, DHCPv6 L2/L3 Relay	15.4.5 Configuring DHCP Relay
	Management information base (MIB) for IPv6 address management	-
IPv6 stack and IPv6 host function	IPv6/IPv4 dual-stack to ensure compatibility of IPv6 and IPv4	-
	Basic IPv6 protocols, including ICMPv6, TCP6, UDP6, and RawIP6	-
	IPv6 Neighbor Discovery (ND) protocol and static configuration of IPv6 neighbors	17.4.5 Configuring IPv6 Neighbor Discovery
	IPv6 PMTU	17.4.3 Configuring PMTU
	IPv6 ping and tracert	• ping ipv6 • tracert ipv6
	IPv6 statistics query and clearance	• display ipv6 statistics • reset ipv6 statistics
IPv6 route	IPv6 static routes	16.5.8 Configuration Example of the IPv6 Static Route
	BGP4+	16.11.13 Configuration Example of BGP4+
IPv6 QoS and security	IPv6 ACL	14.10.6 Configuring Traffic Management Based on ACL Rules
	Anti-MAC spoofing	28.3.3 Configuring MAC Anti-spoofing
	Anti-IPv6 spoofing	27.6.3 Configuring Anti-IP Spoofing
	Anti-denial of service (DoS) attack	26.2.3 Configuring DoS Anti-attack
DAD Proxy and Proxy advertisement for NS	DAD Proxy	ipv6 dad proxy
	Proxy advertisement for neighbor solicitation (NS) on the network	-

Feature	Sub-feature	Configuration Process or Command
	side	
IPv6 Layer 2 Transparent Transmission	Differentiation of service virtual ports based on the IPv6 over Ethernet (IPv6oE) type (0x86DD) and defining of VLANs for service virtual ports	service-port
	Transparent transmission of IPv6 over PPPoE packets	-
	VLAN-based transparent transmission of IPv6 packets	-

NOTE

In this manual:

- For the IPv6 features that are different from IPv4 features, configuration procedures and examples are provided for both IPv6 and IPv4 features.
- For the IPv6 features that are similar with IPv4 features, configuration procedures and examples are not provided for IPv6 features because they are the same as IPv4 features. To configure these IPv6 features, use IPv6 commands and follow the procedures of IPv4 features.

17.4.1 Configuring an IPv6 Address for an Interface

The MA5600T/MA5603T/MA5608T can communicate with other IPv6 equipment only after its interface is configured with an IPv6 address. Before an IPv6 global unicast address or IPv6 link-local address is configured on an interface, the IPv6 packet forwarding function must be enabled on the device.

Context

Each interface can be configured with a maximum of 10 IPv6 global unicast addresses but only one IPv6 link-local address.

- An IPv6 global unicast address is equivalent to an IPv4 public address. It is used for forwarding data across the public network and is necessary for the communication between users. An EUI-64 address has the same function as an IPv6 global unicast address. The difference is that only the network bits need to be specified for the EUI-64 address and the host bits are transformed from the media access control (MAC) addresses of the interface, while a complete 128-bit address needs to be specified for the IPv6 global unicast address.
- The IPv6 link-local address is used in neighbor discovery (ND), and for the communication between nodes on the local link in the stateless address autoconfiguration (SLAAC) process. The packets using the link-local address as the source or destination address are not forwarded to other links.

  The link-local address can be automatically generated or manually configured. It is recommended to automatically generate a link-local address because the link-local address is used to implement communication requirements of protocol and is not directly related to the communication between users.

NOTE
In the SLAAC process, a host uses the prefix information and local interface ID obtained from the received router advertisement (RA) to automatically generate an IPv6 address, rather than using the stateful address autoconfiguration mechanism (DHCPv6).

The MA5600T/MA5603T/MA5608T supports IPv6 address configuration on the virtual local area network (VLAN) interface, METH interface, and loopback interface. This topic uses the VLAN interface as an example.

Procedure

Enable IPv6 packet forwarding capability.

Enabling the IPv6 function on the device and the interface is a prerequisite for configuring IPv6 features. To enable a device to forward IPv6 packets, you must enable the IPv6 capability in both the global config mode and the interface mode.

By default, the IPv6 function is disabled on the device and interface.

1. In global config mode, run the **ipv6** command to enable the IPv6 packet forwarding capability.
2. Run the **interface vlanif** command to enter the VLAN interface mode.
3. In VLAN interface mode, run the **ipv6 enable** command to enable the IPv6 function on the interface.

NOTE
Before configuring other IPv6 features on an interface, you must enable the IPv6 function in interface mode.

Step 1 Configure an IPv6 global unicast address on the interface.

In VLAN interface mode, run the **ipv6 address** or **ipv6 address eui-64** command to configure an IPv6 global unicast address on the interface.

To implement successful communication, you can configure both the EUI-64 address and the IPv6 global unicast address, or configure only one of them. The IPv6 addresses configured on one interface cannot be in the same network segment.

Step 2 Configure an IPv6 link-local address on the interface.

In VLAN interface mode, use either of the following methods to configure the IPv6 link-local address:

- Run the **ipv6 address auto link-local** command to automatically generate a local-link address on the interface.

  After this command is executed, the deletion of the global unicast address does not affect local link communication. If the device only needs to communicate with another device that is directly connected to the device, using the link-local address saves IPv6 global unicast address resources.
- Run the **ipv6 address link-local** command to manually configure a link-local address on the interface. The prefix of the IPv6 address configured by running this command must be FE80::/10.

  If an automatically allocated link-local address exists on the interface, the link-local address will be overwritten by the new link-local address after this command is executed. If possible, avoid changing the link-local address.

If you have not run either of the preceding commands to configure a link-local address for an interface, a link-local address will be automatically allocated to the interface after an IPv6 global unicast address is configured on the interface.

Step 3 Query the address configuration information about an IPv6 interface.

- Run the **display ipv6 interface** command to query the IPv6 interface information.
- Run the **display ipv6 statistics** command to query the IPv6 packet statistics.

----End

Example

To create VLAN 10, set the IPv6 address of VLAN interface 10 to 2000::1/64, and set the generation mode of the link-local address to the automatic generation mode, run the following commands:

```
huawei(config)#vlan 10
huawei(config)#ipv6
huawei(config)#interface vlanif 10
huawei(config-if-vlanif10)#ipv6 enable
huawei(config-if-vlanif10)#ipv6 address 2000::1 64
huawei(config-if-vlanif10)#ipv6 address auto link-local
```

17.4.2 Configuring an IPv6 Address Selection Policy Table

If multiple IPv6 addresses are configured on an interface of the device, the IPv6 address selection policy table can be used to select the source and destination addresses for packets.

Context

Based on their application, IPv6 addresses can be classified into:

- Link-local addresses and global unicast addresses, based on the effective scope of the IPv6 addresses
- Temporary addresses and public network addresses, based on security levels
- Home addresses and care-of addresses, based on the application in the mobile IPv6 domain
- Physical interface addresses and logical interface addresses, based on the interface attributes

The preceding IPv6 addresses can be configured on the same interface of a device. In this case, the device must select a source address or a destination address from multiple addresses on the interface. In addition, if the device supports the IPv4/IPv6 dual-stack, it must also select IPv4 addresses or IPv6 addresses for communication. For example, if a domain name maps both an IPv4 address and an IPv6 address, the device must select an address to respond to the domain name service (DNS) request of the client.

The IPv6 address selection policy table solves the preceding problems. The table defines a group of address selection rules. The source and destination addresses of packets can be specified or notified to the device based on these rules. This table, similar to a routing table, can be queried by using the longest matching rule. The source and destination addresses together determine the selection results.

- The source address is selected based on the *label* parameter. The address whose *label* value is the same as the *label* value of the destination address is selected preferentially as the source address.
- The destination address is selected based on both the *label* and the *precedence* parameters. If the *label* values of the candidate addresses are the same, the address with the largest *precedence* value is selected preferentially as the destination address.

Procedure

In global config mode, run the **ipv6 address-policy** command to configure the source and destination address selection policy.

The system has the default address selection policy entries. These entries are prefixed with ::1, ::, 2002::, FC00::, and ::FFFF:0.0.0.0.

The system supports a maximum of 50 address selection policy entries.

Step 1 Query the IPv6 address selection policy.

Run the **display ipv6 address-policy** command to query the IPv6 address selection policy.

----End

Example

To create the IPv6 address selection policy for the IPv6 address 3::/64, set the precedence to 3, and set the label to 2, run the following commands:

```
huawei(config)#ipv6 address-policy 3:: 64 3 2
huawei(config)#display ipv6 address-policy all
Policy Table :
          Total:6
-------------------------------------------------------------------------------
 Prefix    : ::                                PrefixLength  : 0
 Precedence : 40                                Label        : 1
 Default   : Yes

 Prefix    : ::1                               PrefixLength  : 128
 Precedence : 50                                Label        : 0
 Default   : Yes

 Prefix    : ::FFFF:0.0.0.0                      PrefixLength  : 96
 Precedence : 10                                Label        : 4
 Default   : Yes

 Prefix    : 3::                               PrefixLength  : 64
 Precedence : 3                                 Label        : 2
 Default   : No

 Prefix    : 2002::                             PrefixLength  : 16
 Precedence : 30                                Label        : 2
 Default   : Yes

 Prefix    : FC00::                             PrefixLength  : 7
 Precedence : 20                                Label        : 3
 Default   : Yes
```

```
--------------------------------------------------------------------------------
```

17.4.3 Configuring PMTU

By setting the path maximum transmission unit (PMTU), the device can select a proper maximum transmission unit (MTU) for packet transmission. In this manner, packets do not have to be fragmented during transmission. Employing PMTU reduces the load on intermediate devices, improves network resources utilization, and achieves optimal throughput on the network.

Context

Dynamic PMTU is enabled on a device by default, ensuring that the smallest MTU value is used on all interfaces along the path from the source to the destination node. You can also configure static PMTU to specify the maximum length of a packet that can be forwarded from the source to the destination node. Configuring static PMTU protects devices on a network from jumbo packets. Static PMTU has a higher priority over dynamic PMTU. When both static PMTU and dynamic PMTU are configured, only static PMTU takes effect.

Procedure

Configure static PMTU.

In global config mode, run the **ipv6 pathmtu** command to configure a static PMTU value for a path destined for a specified IPv6 address. By default, the PMTU of the path destined for an IPv6 address is 1500 bytes.

The static PMTU value should be smaller than or equal to the MTU value of every interface on the same path. If the static PMTU value is larger than the MTU value of the interfaces, the system segments packets according to the MTU value. By manually configuring the static PMTU value based on the smallest MTU value of the path over which packets are transmitted, you can achieve a higher packet transmission rate.

Step 1 Configure the aging time of dynamic PMTU entries.

Run the **ipv6 pathmtu age** command to configure the aging time of dynamic PMTU entries. By default, the aging time of dynamic PMTU entries is 10 minutes.

The **ipv6 pathmtu age** command is used to modify the lifetime of dynamic PMTU entries in the buffer. The aging time is invalid on static PMTU entries, because static PMTU entries do not age.

Step 2 Query the PMTU information.

- Run the **display ipv6 pathmtu** command to query the PMTU information.
- Run the **display ipv6 interface** command to query the current MTU value of an IPv6 interface.

----End

Example

To set the PMTU for the IPv6 address 3001::1 to 1300 bytes, and to set the aging time of the dynamic PMTU entries to 40 minute, run the following commands.

```
huawei(config)#ipv6 pathmtu 3001::1 1300
huawei(config)#ipv6 pathmtu age 40
```

17.4.4 Configuring TCP6

Setting TCP6 packet parameters properly helps improve network performance.

Procedure

- Configure TCP6 timers.

  Configuring two TCP6 timers in global config mode helps to control the TCP6 connection time. You are advised to configure the TCP6 timers by following the instructions of technical support engineers.

  a. Run the **tcp ipv6 timer syn-timeout** command to configure the Transfer Control Protocol (TCP) SYN-WAIT timer.

     By default, the SYN-WAIT timer value is 75s.

  b. Run the **tcp ipv6 timer fin-timeout** command to configure the TCP FIN-WAIT timer.

     By default, the FIN-WAIT timer value is 675s.

- Configure the size of the TCP6 sliding window.

  In global config mode, run the **tcp ipv6 window** command to configure the size of the TCP6 sliding window, that is, the sizes of the receiving buffer and transmitting buffer in the TCP6 socket.

  By default, the size of the TCP6 sliding window is 8 KB.

- Query the TCP6 configuration.
  - Run the **display tcp ipv6 statistics** command to query the TCP6 statistics.
  - Run the **display tcp ipv6 status** command to query the TCP6 connection status.
  - Run the **display ipv6 socket** command to query the socket information.

----End

17.4.5 Configuring IPv6 Neighbor Discovery

The IPv6 Neighbor Discovery (ND) protocol provides a set of packets and processes for establishing the relationship between neighboring nodes. The IPv6 ND protocol supports the Address Resolution Protocol (ARP) messages, and the Router Discovery and Redirect messages of the Internet Control Message Protocol (ICMP) of IPv4. In addition, ND supports neighbor reachability detection.

Prerequisite

The IPv6 address has been configured. For details on the configuration method, see 17.4.1 Configuring an IPv6 Address for an Interface.

Context

Most of the ND configurations are implemented based on the interface. The MA5600T/MA5603T/MA5608T supports ND configurations on VLAN Layer 3 interfaces.

Procedure

- Configure the static IPv6 neighbor.

  By configuring a static neighbor, the device can obtain the mapping of the IPv6 address and MAC address of the neighbor. The statically configured neighbor entries will overwrite the dynamically learned neighbor entries and will not age.

  a. In global config mode, run the **interface vlanif** command to enter the VLAN interface mode.

  b. In VLAN interface mode, run the **ipv6 neighbor** command to configure a static IPv6 neighbor.

- Configure the parameters of the Router Advertisement (RA) message.

  The device periodically sends RA messages that contain information such as address prefixes, hop limit value, neighbor reachable time, and message lifetime. The IPv6 node on the local link receives the RA messages and updates its information, such as the IPv6 prefix list and other configuration data, according to the RA messages.

  a. In global config mode, run the **interface vlanif** command to enter the VLAN interface mode.

  b. Run the **undo ipv6 nd ra halt** command to enable RA message advertising.

     - When a device is connected to an IPv6 node, the RA message advertising function needs to be enabled so that the device periodically sends RA messages to the IPv6 node.
     - When a device is not connected to an IPv6 node, the RA message advertising function does not need to be enabled. By default, this function is disabled.

     By default, RA message advertising is disabled on the device.

  c. (Optional) Configure parameters carried in an RA message.

     Perform the following operations as needed:

     - Run the **ipv6 nd ra** command to configure the interval for advertising RA messages.

       By default, the maximum interval is 600 seconds and the minimum interval is 200 seconds. The maximum interval cannot be shorter than the minimum interval. When the maximum interval is shorter than 9 seconds, the minimum interval is adjusted to the same value as the maximum interval.

     - Run the **ipv6 nd ra router-lifetime** command to configure the RA message lifetime.

       By default, the lifetime is 1800 seconds. The lifetime of the message advertised by the device must be longer than or equal to the interval at which the device advertises RA messages.

     - Run the **ipv6 nd ra prefix** command to configure the address prefixes to be advertised in RA messages.

       By default, RA messages contain only the address prefixes specified through the **ipv6 address** command. Run the **ipv6 nd ra prefix** command when you need the device to advertise only the specified prefixes.

     - Configure the default router priority and route information carried in an RA message.

       RA messages that carry the default router priority and route information are advertised over the local link. In this manner, a proper device can be selected to forward messages of a host.

       - Run the **ipv6 nd ra preference** command to configure the default router priority carried in an RA message.

- Run the **ipv6 nd ra route-information** command to configure the route information carried in an RA message.

- Set the stateful autoconfiguration flag bit carried in an RA message.
  - Run the **ipv6 nd autoconfig managed-address-flag** command to set the flag bit carried in an RA message for stateful address autoconfiguration.

    After this flag bit is set, the host uses a stateful address autoconfiguration protocol; otherwise, the host uses a stateless address autoconfiguration protocol.
  - Run the **ipv6 nd autoconfig other-flag** command to set the flag bit for stateful autoconfiguration of other information.

    After this flag bit is set, the host uses a stateful autoconfiguration protocol for configuring information other than the address.
- Run the **ipv6 nd ns retrans-timer** command to configure the interval for detecting neighbor reachability, that is, the neighbor solicitation (NS) message retransmission timer.

  Frequently sending NS packets helps to determine the neighbor reachability but also affects the device performance. Therefore, you are not advised to set the interval to a small value. The default interval, 1000 milliseconds, is recommended.
- Run the **ipv6 nd nud reachable-time** command to configure the neighbor reachable time specified in an RA message.

  The device detects neighbor reachability by using the neighbor unreachability detection (NUD) mechanism. The neighbor reachable time configured through the **ipv6 nd nud reachable-time** command is the interval for running a NUD detection. A smaller neighbor reachable time set on a device means that the device can probe the neighbor reachability more quickly but, meanwhile, more network bandwidth and CPU resources will be consumed. Therefore, you are not advised to set the neighbor reachable time to a short interval. The default interval, 30000 milliseconds, is recommended.
- Configure the hop limit value for the router.

  Run the **quit** command to quit the VLAN interface mode. In global config mode, run the **ipv6 nd hop-limit** command to configure the hop limit value for the router, that is, the maximum number of hops for the IPv6 unicast packets initiated by the router. The hop limit value for the router is the same as the hop limit value specified in the RA message. The default value is 64.

- Configure duplicate address detection (DAD).

  DAD is used to check whether an IPv6 address is available. When a node is configured with an IPv6 address, it immediately sends an NS message to check whether this address is already used by other neighboring nodes.

  a. In the global config mode, run the **interface vlanif** command to enter the VLAN interface mode.

  b. Run the **ipv6 nd dad attempts** command to configure the number of DAD attempts, that is, number of attempts to send NS messages. The default value is 1.

  c. Run the **ipv6 nd ns retrans-timer** command to configure the interval of DAD, that is, the timer for retransmitting NS messages. The default interval, 1000 milliseconds, is recommended.

- Configure the ND proxy.

  a. In the global config mode, run the **interface vlanif** command to enter the VLAN interface mode.

  - Run the **nd proxy enable** command to enable the ND proxy on the interface of a non-super VLAN, and enable the ND proxy between the sub-VLANs under a super VLAN interface.
  - Run the **nd proxy subvlan** *vlanid* **[to** *end-vlanid* **]** command to enable the ND proxy in the sub-VLANs under a super VLAN interface.
- Query the IPv6 neighbor information.
  - Run the **display ipv6 neighbors** command to query the IPv6 neighbor information.
  - Run the **display ipv6 interface** command to query the IPv6 interface information.
  - Run the **display ipv6 prefix** command to query the IPv6 prefixes in the RA message sent from the IPv6 interface.
  - Run the **display ipv6 route-information** command to query the route information in the RA message sent from the IPv6 interface.

----End

Example

To configure the function of automatically generating a link-local unicast address on VLAN interface 10, set the prefix to be advertised by the local unicast address of site EUI-64 and by the RA message to 3000::/64, set both the valid lifetime and the preferred lifetime of the prefix to 1000s, and enable RA message advertising (so that the host can automatically obtain the address prefix in the RA message), run the following commands:

```
huawei(config)#vlan 10
huawei(config)#ipv6
huawei(config)#interface vlanif 10
huawei(config-if-vlanif10)#ipv6 enable
huawei(config-if-vlanif10)#ipv6 address auto link-local
huawei(config-if-vlanif10)#ipv6 address 3000::/64 eui-64
huawei(config-if-vlanif10)#ipv6 nd ra prefix 3000::/64 1000 1000
huawei(config-if-vlanif10)#undo ipv6 nd ra halt
```

17.5 Reference Standards and Protocols

The following table lists the reference standards and protocols of the IPv6 feature.

Standard/Protocol	Description
RFC1887	An Architecture for IPv6 Unicast Address Allocation
RFC1981	Path MTU Discovery for IP version 6
RFC2375	IPv6 Multicast Address Assignments
RFC2460	Version 6 of the Internet Protocol (IPv6), also sometimes referred to as IP Next Generation or IPng.
RFC2461/RFC4861	Neighbor Discovery for IP Version 6 (IPv6)
RFC2462/RFC4862	IPv6 Stateless Address Auto configuration
RFC2463/RFC4443	Internet Control Message Protocol (ICMPv6) for the Internet Protocol Version 6 (IPv6) Specification

Standard/Protocol	Description
RFC2464	Transmission of IPv6 Packets over Ethernet Networks
RFC2466	Management Information Base for IP Version 6 ICMPv6 Group
RFC2526	Reserved IPv6 Subnet Anycast Addresses
RFC2711	IPv6 Router Alert Option
RFC2893	Transition Mechanisms for IPv6 Hosts and Routers
RFC3633	IPv6 Prefix Options for Dynamic Host Configuration Protocol(DHCP) version 6
RFC3736	Stateless Dynamic Host Configuration Protocol (DHCP) Service for IPv6
RFC3849	IPv6 Address Prefix Reserved for Documentation
RFC4001	Textual Conventions for Internet Network Addresses
RFC4007	IPv6 Scoped Address Architecture
RFC4191	Default Router Preferences and More-Specific Routes
RFC4193	Unique Local IPv6 Unicast Addresses
RFC4213	Basic Transition Mechanisms for IPv6 Hosts and Routers
RFC4214	Intra-Site Automatic Tunnel Addressing Protocol(ISATAP)
RFC4429	Duplicate Address Detection
RFC4282	A Model of IPv6/IPv4 Dual Stack Internet Access Service
RFC2373/RFC3513/RFC4291	Internet Protocol Version 6 (IPv6) Addressing Architecture
RFC4862/RFC5006	Router Advertisement (RA) filtering
RFC6221	DHCPv6 LDRA

18 Multicast

About This Chapter

Multicast is a communication mode in which data is transmitted to multiple recipients at the same time.

18.1 Introduction to Multicast

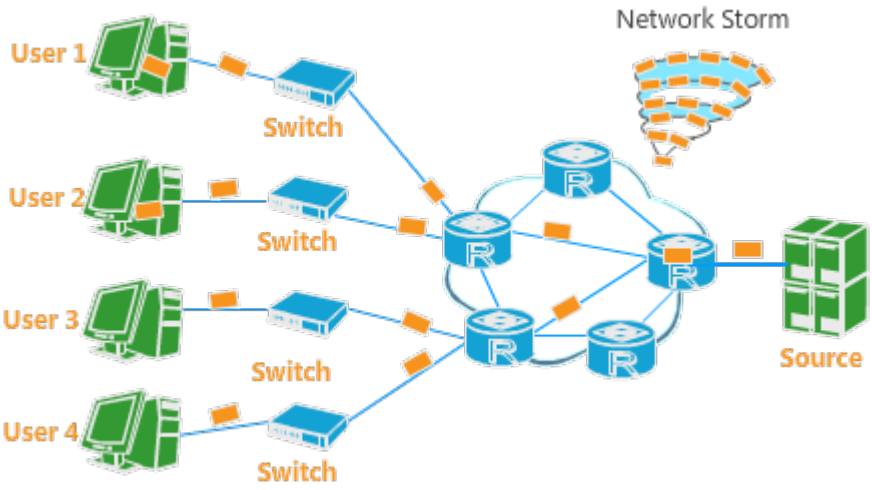
Why Is Multicast Required

Traditional IP communications are implemented in unicast or broadcast mode.

- Unicast: a P2P transmission mechanism. Unicast involves only one information sender and one information recipient.
- Broadcast: a point-to-all-point transmission mechanism. Broadcast involves only one information sender and all the reachable information recipients in a LAN.

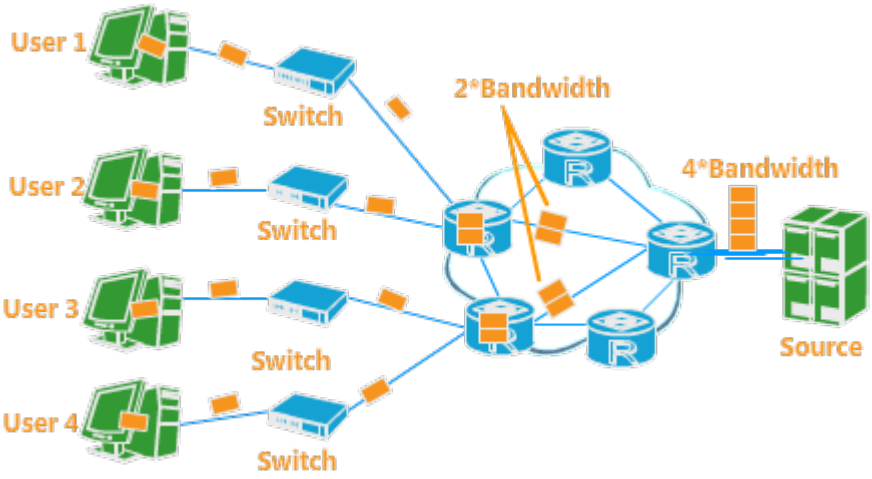
In broadcast IP communications, the source sends only one copy of packets to the broadcast address of the network. Then, the network copies these packets and sends them to all network segments for all routers and users, regardless of whether these routers and users require such data. In broadcast mode, if only a few users require such packets, network utilization is low and bandwidths are unnecessarily used. The users who do not require such data are adversely affected. In addition, a severe broadcast storm may occur due to routing loopback.

Figure 18-1 Diagram of broadcast implementation



In unicast IP communications, the source sends a separate copy of packets to each recipient. Therefore, each recipient requires a separate data channel. In unicast mode, the amount of data that needs to be transmitted on the network is determined based on the number of users requiring such data. If the number of users requiring the same data is large, multiple same data flows must be set up on the network. In this case, network bandwidth may become insufficient, degrading network transmission quality. Therefore, the unicast mode cannot be used for transmitting large amount of data for a large number of users.

Figure 18-2 Diagram of unicast implementation

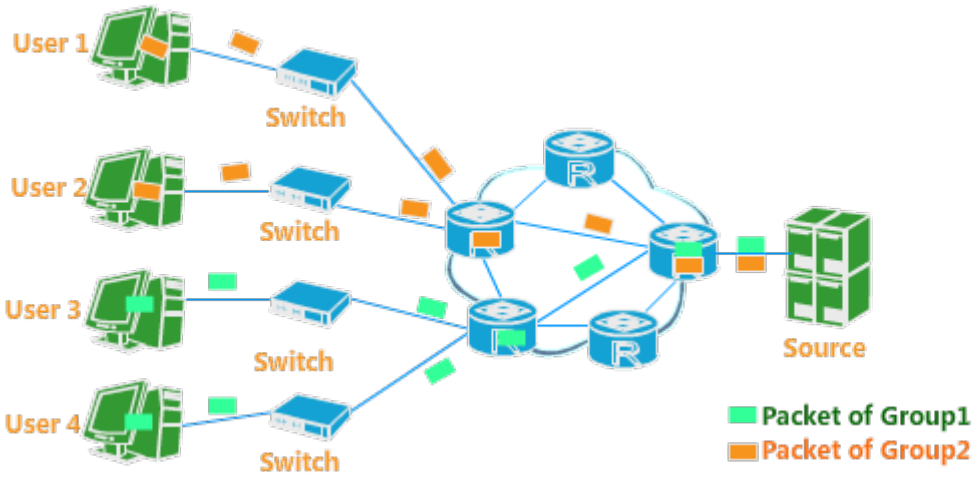


Traditional unicast and broadcast communications cannot meet P2MP service requirements. The IP network development supports more network applications. In addition, some large amount of data needs to be sent in P2MP mode. In this case, multicast is introduced to meet these service requirements.

What Is Multicast

Multicast is a communication mode in which one information sender sends one copy of IP data package to a group of information recipients on an IP network. Other hosts on the network cannot receive this data package.

Figure 18-3 Diagram of multicast implementation



Comparison between multicast and unicast, and multicast and broadcast:

- In the multicast mode, a single data stream is sent to a group of users at the same time. Only one copy of the same multicast data stream exists on each link. Compared with the unicast mode, in the multicast mode, the increase of users does not immediately increase the load of the network. Therefore, the server and the CPU can deal with a lighter load, reducing desired network bandwidths.
- Multicast messages can be sent across different network segments and will not be received by users who are not interested in the messages. Compared with the broadcast mode, the multicast mode achieves a longer information transmission distance and ensures that information is transmitted to only interested recipients. Hence, information security can be guaranteed.

The preceding comparisons show that multicast effectively resolves the problem of P2MP transmission and implements efficient P2MP data transmission in IP networks.

Multicast Applications

The multicast technology enables the device to provide value-added services, including live broadcast, IPTV, distance learning, TeleMedicine, network radio, live radio conference, and online game.

18.2 Basic Multicast Concepts

Basic Concepts

The following section provides an example of watching the program of a TV channel to aid the understanding of relevant concepts of IP multicast.

- The multicast group is an agreement between the sender and the recipient. For example, a TV channel can be regarded as a multicast group.
- The TV station is the multicast source and it sends data to a certain TV channel.
- The STB is a receiving host. When the user chooses to watch the program of a channel, this action can be regarded as the host joining a multicast group. Then, the TV set displays the program of the TV channel to the user, which means that the host has received the data sent to this multicast group.
- The user can turn on or turn off the STB or switch between channels any time, which means that the host can join or leave a multicast group dynamically.

Figure 18-4 Multicast concept diagram

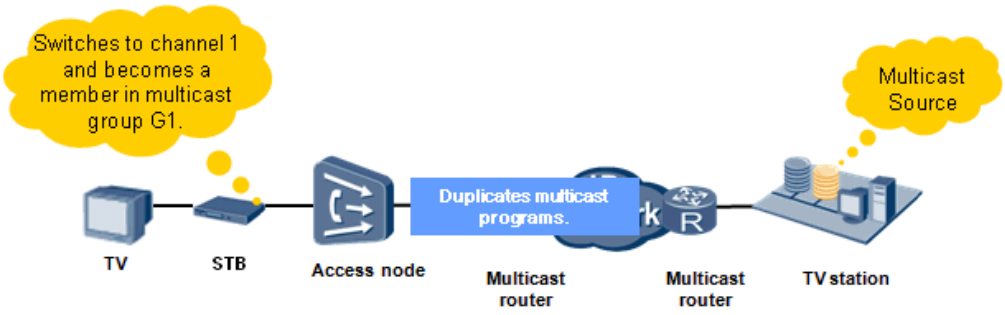
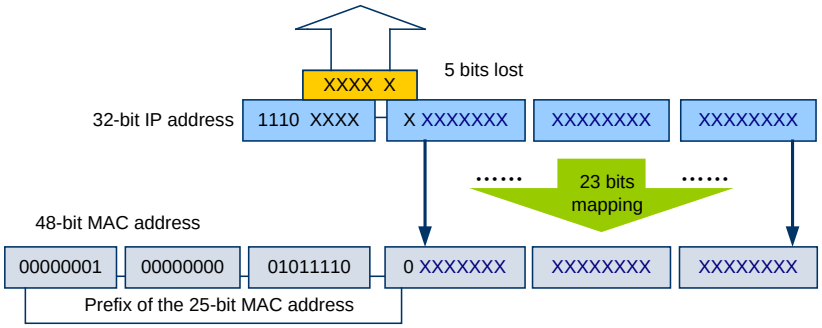


Table 18-1 Table of Basic Concepts

Concept	Description
Multicast group	A multicast group is identified by a multicast IP address. Any host (or any other receiving device) joining a multicast group becomes a member of the group. The group member can identify and receive the IP messages destined to the multicast IP address.
Multicast source	A signal source sending IP messages destined to a multicast address is called a multicast source. • A multicast source can send data to multiple multicast groups at the same time. • Multiple multicast sources can send data to a multicast group at the same time.
Multicast group member	The members of a multicast group are dynamic. Hosts in a network can join or leave a multicast group any time. Multicast group members may be widely dispersed across the network. A multicast source is usually not a data recipient at the same time and is not the member of a multicast group.
Multicast duplication	Multicast duplication is a capability with which the network device duplicates a multicast message from an ingress port into multiple copies and sends them to multiple egress ports. To ensure effective transmission of multitudes of data, this function can be implemented only by hardware.
Multicast	To enable the communication between a multicast source and its members, a network-layer multicast address must be available, which is

Concept	Description
address	the multicast IP address. In addition, a technology must also be available for mapping the multicast IP address to a link-layer multicast MAC address. The following part of this section will describe the two types of multicast address. • Multicast IP address As specified by Internet Assigned Numbers Authority (IANA), multicast messages use class-D IP addresses (224.0.0.0-239.255.255.255) as their destination addresses, and the class-D IP addresses must not appear in the source IP address field of the IP messages. For details of the MAC address range and description, see Table 18-2. Multicast addresses are not allocated to the receiving device or the multicast source device for identifying their network location. In the case of the multicast source device, the allocated multicast address is used for generating and carrying multicast data; in the case of the receiving device, the multicast address is used for distinguishing multicast data. In an actual multicast application, the multicast address does not need to be manually input. For example, in the live TV service, which is a common application, a menu interface is provided. When the user orders a program using a remote controller, the application software will automatically obtain the multicast IP address corresponding to the program. • Ethernet multicast MAC address When IP messages are unicast over an Ethernet, the destination MAC addresses used are the MAC addresses of recipients. However, in the transmission of multicast messages, the transmission destination is no longer a specific recipient. Instead, it is a group with uncertain members. In this case, the multicast MAC address is used. Specified by IANA, the most significant 25 bits of a multicast MAC address are 0x01005e, and the least significant 23 bits of the MAC address are the least significant 23 bits of the multicast IP address. The following figure shows the mapping. **Figure 18-5** Mapping between multicast MAC address and multicast IP address  The first four bits of the multicast IP address are 1110, which stands for the multicast ID, and in the last 28 bits, only 23 bits are mapped to the

Concept	Description
	MAC address. Therefore, five bits of information in the IP address is lost. The direct result is that 32 multicast IP addresses are mapped to the same MAC address.
Multicast router	Multicast routers support multicast. A multicast router supports the following functions: • Manages group members in the network segment close to user hosts. • Supports multicast routing for forwarding multicast packets. • Functions as a group member.

Table 18-2 Multicast addresses and meanings

MAC Address Range	Description
224.0.0.0-224.0.0.255	Permanent group addresses reserved for routing protocols
224.0.1.0-238.255.255.255	User multicast addresses
239.0.0.0-239.255.255.255	Local management group addresses (private addresses) **NOTE** This MAC address range limits multicast to be implemented within a specified multicast domain so that the MAC addresses in different domains can be duplicated.

Basic Managed Objects

Table 18-3 Basic managed objects

Managed Object	Description
Multicast VLAN	A multicast VLAN (also called an MVLAN) refers to the VLAN tag carried by multicast data. MVLANs are usually divided based on ISP. By the implementation of the forwarding plane, control plane, and management plane based on VLAN instance, multicast services are provisioned to the users of the same device, allowing the users not to be interfered by each other. Except the super VLAN, the VLAN of any attribute or any type configured on the device can serve as an MVLAN. For details on the MVLAN, see "Multi-instance Multicast".
Multicast program	A multicast program can be regarded as a multicast group. Its basic attribute is the multicast IP address. The device can manage a multicast program at a finer grain, such as by rights control and CAC. According to whether the attributes (such as the multicast IP address) of

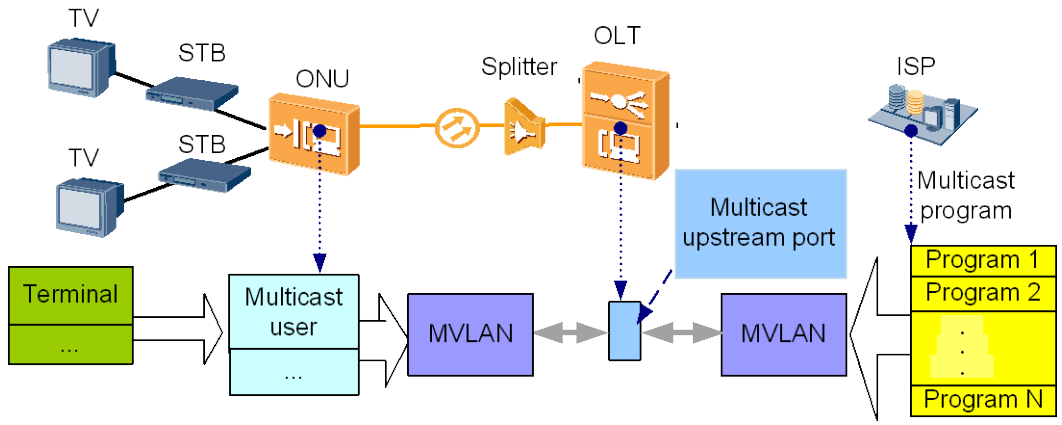
Managed Object	Description
	each program are configured before the service is provisioned, multicast programs can be classified into two types: pre-configured programs and dynamic programs. Determine whether to support the programs of the two types based on the configured MVLAN program mapping mode. Table 18-4 shows the supporting status for pre-configured programs and dynamic programs. For details about dynamic programs, see Dynamic Programs.
Multicast uplink port	A multicast uplink port is one through which a multicast source connects to the device, and is also a port through which an upper-layer multicast router connects to the device. According to their dependency on the link-layer loop protocol, multicast uplink ports can be classified into two types: manually configured (static) uplink ports and dynamic uplink ports. For details on dynamic uplink ports, see Dual-homing of Upstream Ports.
Multicast user	A multicast user is a multicast data recipient. A service stream must be configured for the multicast user for carrying multicast control messages in the upstream direction (the device can distinguish the user by traffic classification). Therefore, a multicast user corresponds to a unique terminal or service subscriber. In addition, an MVLAN must be specified for the multicast user to indicate to which ISP the service subscriber belongs.

Table 18-4 Program type supported

Program Type	Enabled	Disabled
Pre-configured programs	Supported	Supported
Dynamic programs	Not supported	Supported

The following figure shows the relationships between the basic managed objects.

Figure 18-6 Multicast managed objects



NOTE

- As a device placed at users' home, an ONT does not support multiple multicast users. Although the ONT is for only one multicast user, it can still connect to multiple STBs.
- Traffic streams with the QinQ attribute do not support multicast users.
- Traffic streams that classify traffic by double VLANs do not support multicast users.

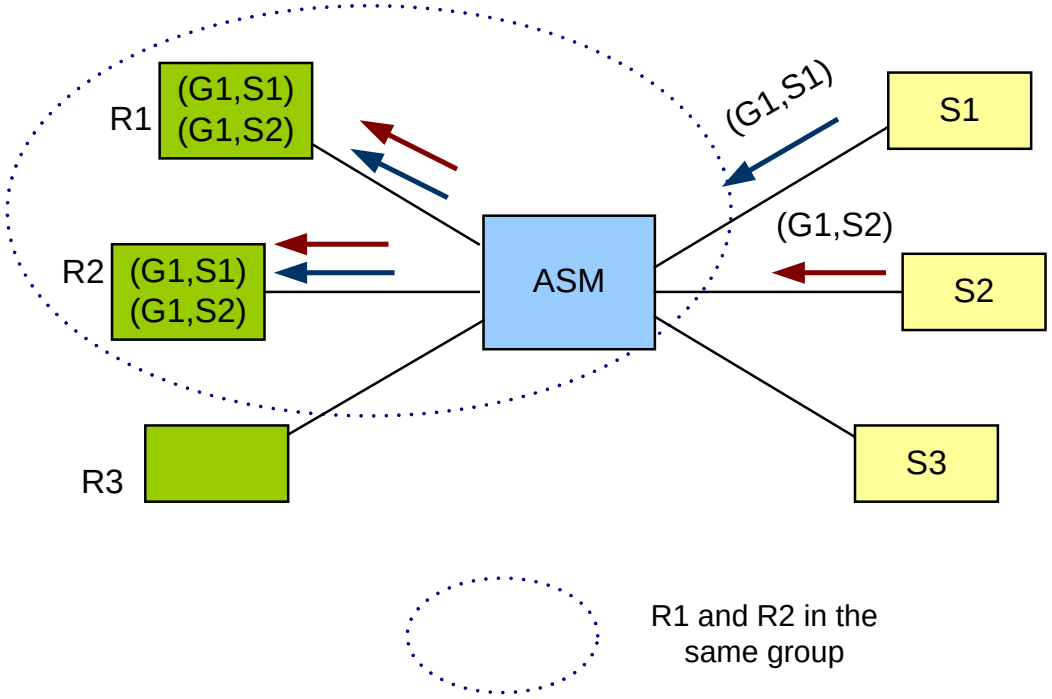
18.3 Multicast Model

According to the multicast source control level, multicast has three models:

ASM

Any-source multicast (ASM) is defined in RFC 1112. In this model, a recipient, by joining a group identified by the multicast address, can receive data sent to the group. A recipient can join or leave a group at any time, and the recipient location or quantity is not limited. In addition, any sender can serve as the multicast source to send data to the group. Therefore, this model is applicable to the multipoint-to-multipoint (MP2MP) multicast application.

Figure 18-7 ASM network model



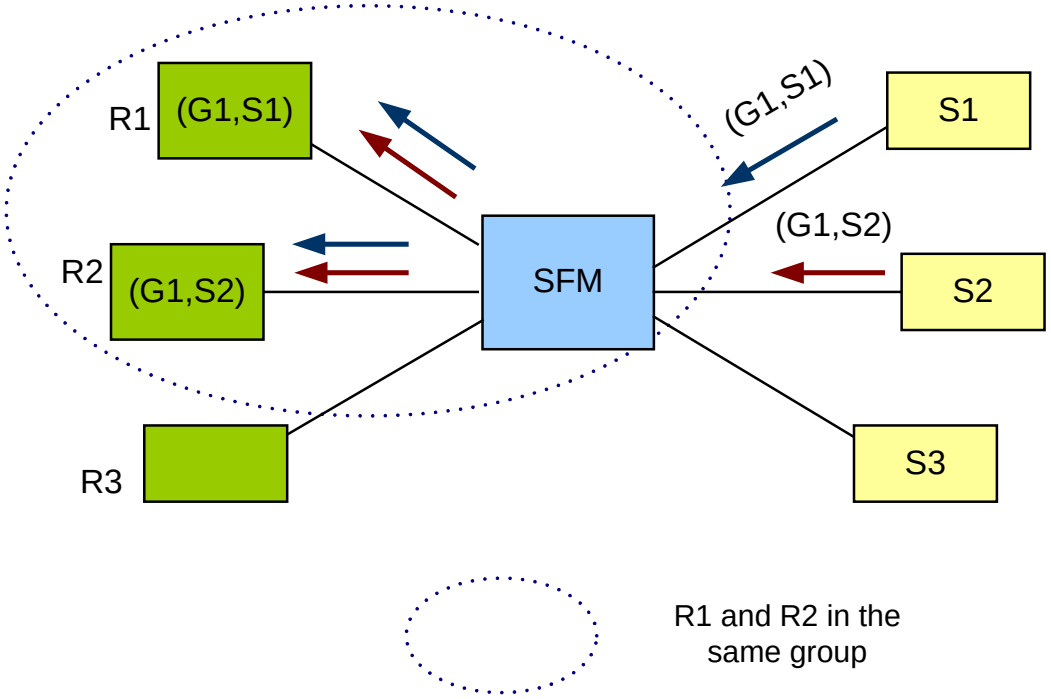
SFM

As an extension of ASM, source-filtered multicast (SFM) extends the source filtering function of the upper-layer protocol module. That is, in the SFM model, whether the multicast data of specified multicast source(s) is allowed to pass can be controlled. Viewed from recipients, SFM and ASM are different; but viewed from senders, they are the same. Therefore, SFM is the same as ASM in terms of network interoperability.

NOTE

The SFM is basically an ASM with a multicast source filtering policy. The basic principles and configurations of ASM and SFM are the same. In this manual, both SFM and ASM are called ASM.

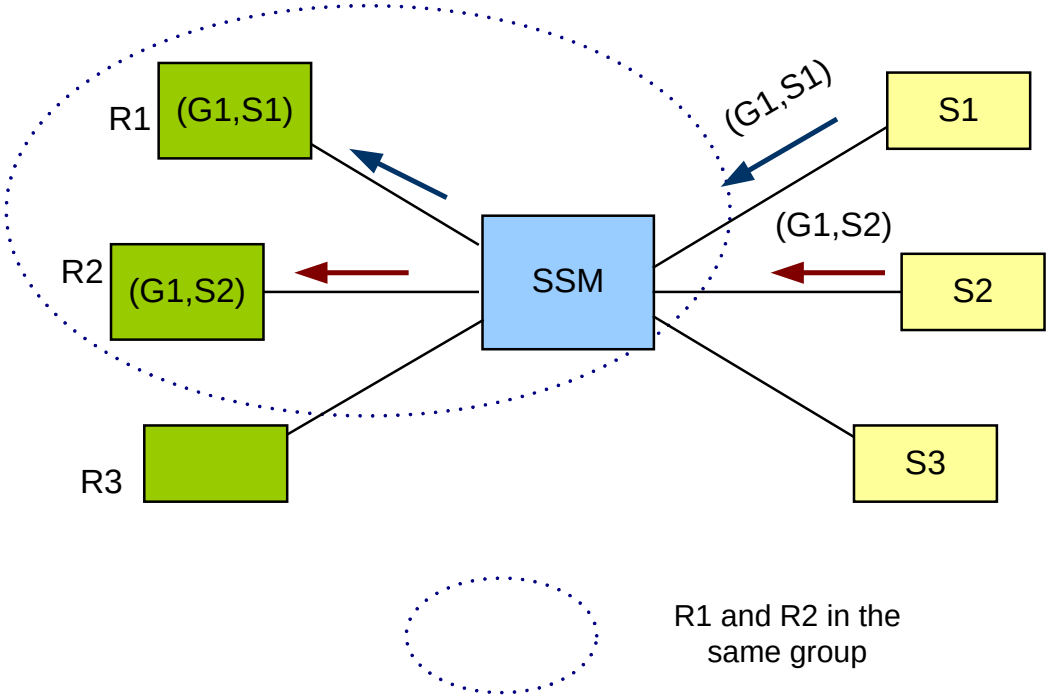
Figure 18-8 SFM network model



SSM

Source-specific multicast (SSM) is defined in RFC 4607. In this model, a recipient joins a channel by specifying the multicast source and group and receives data sent to the group from a specific multicast source. The recipient quantity is not limited. In addition, only the specific sender can serve as the multicast source to send data to the channel. Therefore, this model is applicable to the point-to-multipoint (P2MP) multicast application.

Figure 18-9 SSM network model



The following table lists the protocols that support ASM/SSM.

Multicast Model	Typical Protocol Combination of Devices in the Network			
	STB	AN	Router	Inter-domain router
ASM	IGMPv2	IGMPv2	PIM-SM	MSDP/MBGP
SSM	IGMPv3	IGMPv3	PIM-SSM	MBGP

Based on the preceding multicast models, the OLT supports three group filtering modes: ASMSSM, ASM ONLY, and SSM ONLY.

NOTE
Only the OLT supports group filtering mode.

In different group filter modes, for the differences in management plane and control plane, see Multi-instance Multicast; for the differences in the forwarding plane, see Forwarding Framework on the Device.

Follow the rules below to select a proper filter mode:

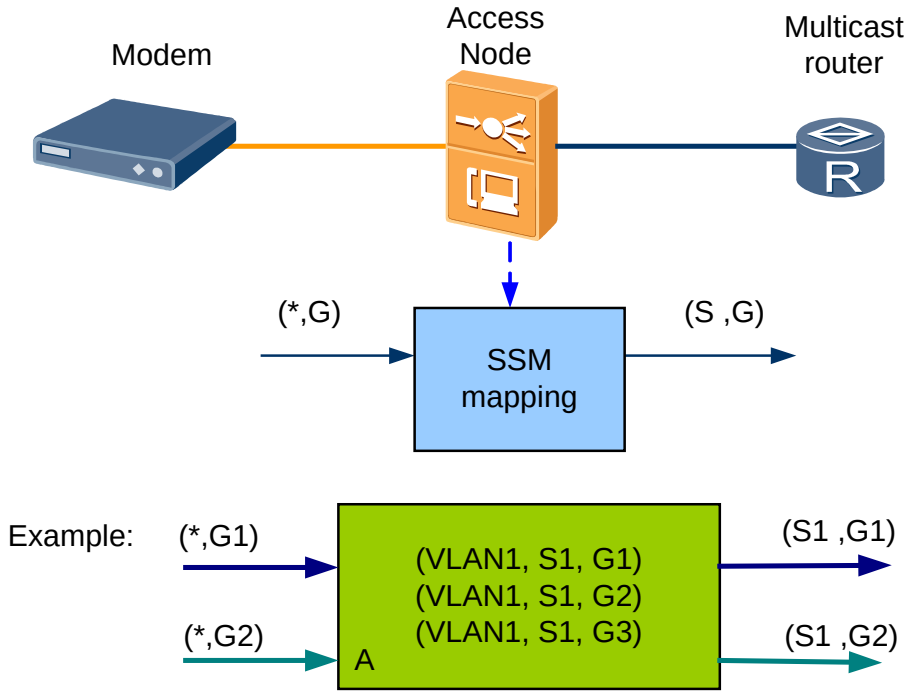
- For compatibility with the original VLAN+GMAC, select ASMSSM.
- For strict ASM or SSM message filtering, select ASM ONLY or SSM ONLY.
- For maximum compatibility of various STBs in a home network, select ASMSSM.
- For SIP+GIP forwarding (that is, posing no restrictions on SIP and GIP planning among different ISPs), select SSM ONLY. However, the entire network needs to support IGMPv3 for implementing SSM ONLY.

When the group filter mode is ASMSSM, even if the user side does not support IGMPv3, carriers can implement SSM network on the network side. The device, with its SSM mapping function (can be supported automatically), can help implement the SSM networking on the network side even if the user-side device does not support IGMPv3. With the SSM mapping function, the device maps the received (*, G) message to an (S, G) message according to the unique multicast program triplet, as shown in the following figure.

NOTE

- A multicast user cannot belong to multiple MVLANs at the same time.
- Dynamic programs do not support SSM mapping.

Figure 18-10 SSM mapping



The following table describes whether the devices support the SSM and ASM modes.

User Side	Network Side	Supported or Not
SSM	SSM	Supported
ASM	ASM	Supported
ASM	SSM	Supported
SSM	ASM	Not supported

NOTE

Currently the MxU does not support the PIM-SSM.

When the group filter mode is ASM ONLY, SSM mapping is not supported because ASM packets are dropped on the user side. For SSM ONLY, it is unnecessary to support SSM mapping.

18.4 Implementation Principles of Multicast

NOTE

- The MA5600T/MA5603T/MA5608T can function as an OLT or DSLAM. In the remainder of this document, [OLT] and [DSLAM] indicate multicast implementation principles on the MA5600T/MA5603T/MA5608T that function as an OLT and DSLAM, respectively.
- [DSLAM/ONU] indicates the implementation principles of the ONU device is the same as the MA5600T/MA5603T/MA5608T that function as a DSLAM.
- [ONU] indicates the implementation principles of the ONU device.

18.4.1 IGMP

IGMP Protocol

Internet Group Management Protocol (IGMP) is used to set up and maintain relationships between multicast members in a multicast group between hosts and the router connected to these hosts.

- Multicast members dynamically add to or leave a multicast group on the host.
- The relationships between group members are maintained and managed on the router. In addition, the IGMP protocol supports data exchanging with the upper-layer multicast routing protocol.

IGMP involves three versions: IGMPv1, IGMPv2, and IGMPv3. A new IGMP version is compatible with an earlier version. IGMPv1 is rarely supported. Therefore, according to TR101 requirements, Huawei hosts do not support IGMPv1 and simply discard IGMPv1 packets.

IGMPv1

IGMPv1 is defined in RFC 1112, defining the basic process of group member query and report. IGMPv1 is seldom used. Therefore, IGMPv1 is not described in the remainder of this document.

IGMPv2

IGMPv2 is defined in RFC 2236, supporting fast leave of group members compared with IGMPv1.

Role	Message Type	Description
Router	General query	A router periodically sends this message to maintain the requirements posed by all hosts connected to the router in all multicast groups. The router detects an accidentally offline host by an aging mechanism.
	Group-specific query	A router sends this message to check whether a multicast group is still required by any host. The router usually sends this message when receiving a leave message.
Host	Report	The report message is used by a host for actively joining a multicast group or for responding to a general query or a group-specific query.
	Leave	The leave message is used by a host for actively

Role	Message Type	Description
		informing a router that the host no longer needs a multicast group.

Group Member Relationship Maintenance

Prerequisite: If a network segment contains multiple multicast routers, IGMPv2 enables the routers to use a querier election mechanism to elect a querier.

1. The querier periodically sends general query messages to group members to query group member relationships.
2. The hosts in the network segment make different responses to the querier after receiving the general query messages.
   - The hosts out of the multicast group do not respond to the querier.
   - Group members locally start a timer.
     - If a group member does not monitor a report message responded by other group members when the timer times out, this group member sends a report packet to respond to the querier.
     - If a group member monitors a report message responded by other group members before the timer times out, this group member does not send a report packet to respond to the querier (suppresses its response packet).
3. After receiving the report messages, the querier determines that this network segment contains group members and generates a multicast forwarding entry. When receiving data for the multicast group, the router forwards the data to the group members.

Group Member Join-in

1. If host A added to multicast group G1 monitors no data for G1 in the network segment, host A immediately sends a report message of G1 to the querier, without waiting for the receiving of a general query message sent by the querier.
2. After receiving the report message, the router in the network segment determines that this network segment contains G1 group members and adds the downlink port for G1 to multicast routing entry (*,G1). When receiving data for G1, the router forwards the data to the group members.

Group Member Leaving

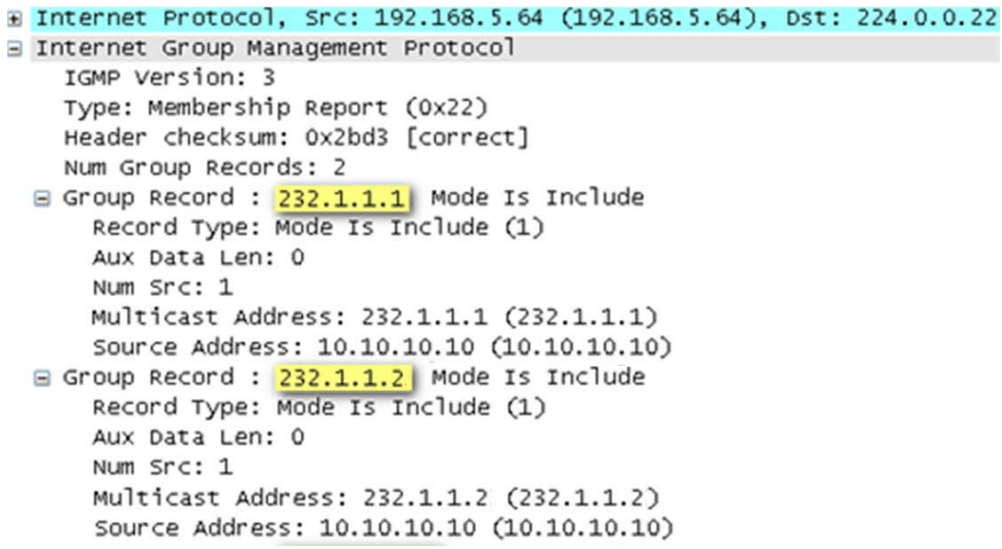
1. When host A exits from G1, it sends a leave message to the querier.
2. After receiving the leave message, the querier sends a group-specific query message to group members to check whether all group members in G1 have left this group.
   - If a group member sends a report message to the querier, this network segment still contains G1 group members. Therefore, the querier does not delete the downlink port from multicast entry (*,G1).
   - If no group member sends a report message to the querier, the querier deletes the downlink port from multicast entry (*,G1).

IGMPv3

IGMPv3 is defined in RFC 3376. Compared with IGMPv2 (RFC 2236), IGMPv3 has the following improvements:

- Batch report. The destination IP address of report messages is always filled in as 224.0.0.22. In addition, the IGMP payload can carry multiple group records, reducing the number of report messages between devices. As shown in the following figure, the IGMP message captured by a packet capture tool carries the information about two groups 232.1.1.1 and 232.1.1.2 With IGMPv2 messages, the destination IP address must be filled in as the corresponding group IP address. Hence, one IGMPv2 message cannot carry the information about multiple groups.

  Figure 18-11 Example of an IGMPv3 report message



- Longer maximum response time for a query message. In IGMPv3, the maximum response time for the query message is extended from 25.5s (IGMPv2) to 3174.4s. Therefore, IGMPv3 is applicable to large-scale networks.
- Source filter. With the source filter function, the host can receive or not receive the multicast data carrying the IP address of a specified multicast source. This function enables the device to better implement SSM and support the multiple-ISP scenario. IGMPv2 supports only ASM. The following uses different types of messages to explain the implementation of source filter.
  - Query messages

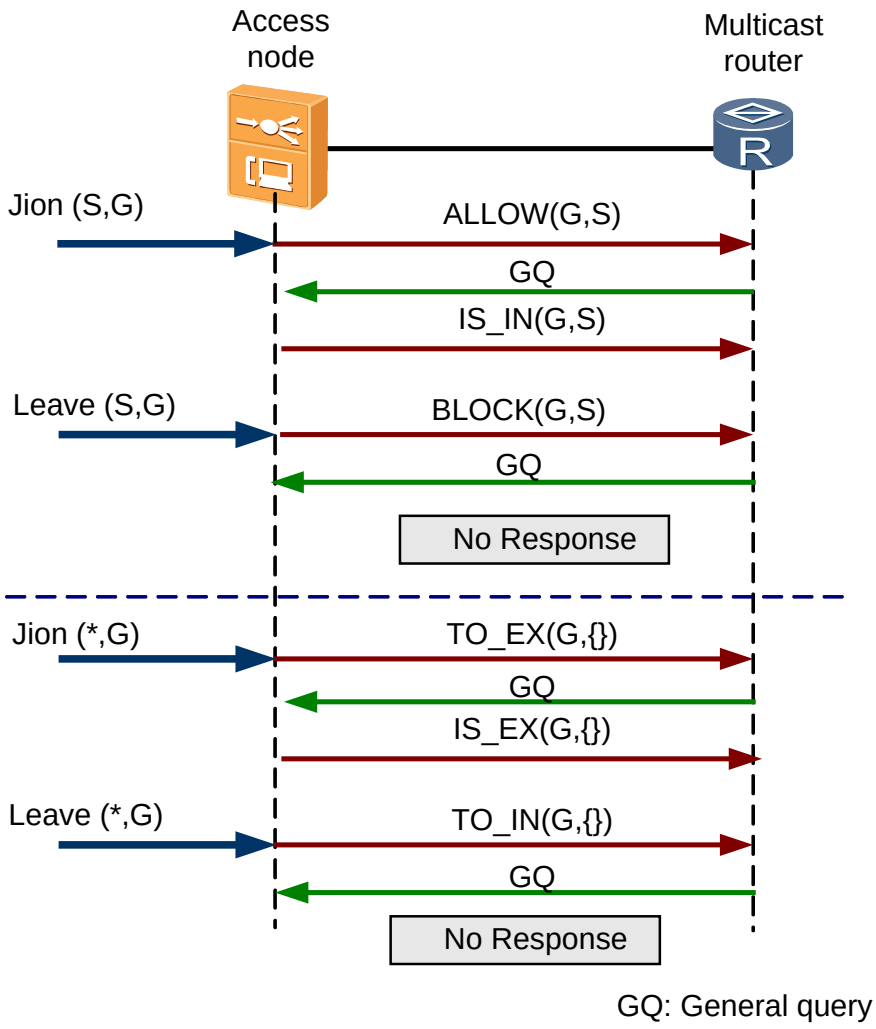
General query	The device sends this message to learn the reception status of an interface to "all" multicast groups. This is similar to the general query of IGMPv2.
Group-specific query	The device sends this message to learn the reception status of an interface to the multicast group with a specific address. This is similar to the group-specific query of IGMPv2.
Group-and-source-specific query	The device sends this message to learn the reception status of an interface to the multicast group with a specific group address and source address. This is a new message of IGMPv3.

  - Report messages

IS_IN(G, S)	Reports the status. Indicates that the current mode of the group is the INCLUDE mode. This message is triggered when the device receives a query message. The source address list contains the source address S of the group.
TO_IN(G, S)	Changes the filter mode of the multicast group to the INCLUDE mode. The source address list contains a new source address S. TO_IN(G, {}) indicates leaving all sources of G and this message in this case is the same as the IGMPv2 leave message.
ALLOW(G, S)	Changes the source address list. This message is triggered when the source address changes. The source address contained in the record is the source address S that the system wishes to join.
BLOCK(G, S)	Changes the source address list. This message is triggered when the source address changes. The source address contained in the record is the source address S that the system does not wish to join.
IS_EX(G, S)	Reports the status. Indicates that the current mode of the group is the EXCLUDE mode. This message is triggered when the device receives a query message. The source address list contains the source address S that the group does not wish to join. IS_EX(G, {}) indicates that the device is interested in all sources of G and this message in this case is the same as the IGMPv2 join message. The device does not support the IS_EX message that contains an empty S.
TO_EX(G, S)	Changes the filter mode of the multicast group to the EXCLUDE mode. The source address list contains a new source address S that the device does not wish to join. TO_EX(G, {}) indicates joining all sources of G and this message in this case is the same as the IGMPv2 join message. The device supports the TO_EX message that contains an empty S.

The following figure shows an example of the report message application.

Figure 18-12 Program ordering behavior converted into IGMPv3 messages



IGMP Version Compatibility

The compatibility policies of the IGMP version on access devices distinguish between the network side and the user side.

The IGMP version on the network side is configured based on MVLAN. As shown in the following table, according to the IGMP version on the multicast router, the IGMP version on the device should be set to the recommended version to avoid incompatibility. Incompatibility may cause packet loss.

Multicast Router	MVLAN on Access Device	Interoperation Result
v1	v2/v3	Incompatible
v2	v2 (recommended)	Normal
v3	v2	Normal. The device can response the IGMPv3 query

Multicast Router	MVLAN on Access Device	Interoperation Result
	(recommended)	messages, and the multicast router can process the IGMPv2 join and leave messages.
v2	v3	The multicast router does not process IGMPv3 messages. Interoperation is normal only after the IGMP version on the device is downgraded to v2. Before the downgrade, packet loss may occur.
v3	v3 (recommended)	Normal

The IGMP version on the user side can be configured based on multicast users. As shown in the following table, according to the IGMP version on the terminal, the IGMP version on the device should be set to the recommended version to avoid incompatibility. Incompatibility may cause packet loss.

Terminal	Multicast User on Access Device	Interoperation Result
v1	v2/v3/v3-forced	Incompatible
v2	v2 (recommended)	Normal
v3	v2	The device does not process IGMPv3 messages. Interoperation is normal only after the IGMP version on the terminal is downgraded to v2 (the terminal can be downgraded by enabling the function of periodically sending query messages to offline users). Before the downgrade, packet loss may occur.
v2	v3 (recommended)	The terminal does not process IGMPv3 messages. Interoperation is normal only after the IGMP version on the device is downgraded to v2. Even after downgraded to IGMPv2, the device can still identify the IGMPv3 messages sent from other terminals. This ensures greater compatibility of the device. In normal application scenarios, the terminal is usually the active initiating party. The IGMP version on the device can be seamlessly downgraded without packet loss.
v3	v3 (recommended)	Normal
v2	v3-forced	Incompatible
v3	v3-forced	The device drops IGMPv2 messages without processing them. Therefore, the device will not be downgraded to v2 but stays in v3.

IGMP Mode

MVLAN-based IGMP modes include IGMP proxy and IGMP snooping.

IGMP Proxy

IGMP proxy is a mode in which the device in a tree topology does not set up a route to forward multicast messages, but only acts as a proxy for multicast protocol messages. Details are as follows:

- From the perspective of a terminal, the device serves as a multicast router that implements the functions of the router in the IGMP protocol. Specifically, the device consistently functions as an IGMP querier (not supporting querier election for security concerns) on the user-side network. The device receives and terminates the join and leave messages of all multicast users, and duplicates the multicast program to only the interested multicast users according to the maintained group membership table.

Table 18-5 Structure of the group membership table-OLT

Group Filter Mode	Index	Online Member
ASMSSM ASM ONLY	VLAN+GIP	Multicast user list (such as multicast user 1 and multicast user 2)
SSM ONLY	VLAN+GIP+SIP	Multicast user list (such as multicast user 1 and multicast user 2)

- From the perspective of a multicast router, the device serves as a multicast group member that implements the functions of the host in the IGMP protocol. According to the changes (addition or deletion) of the record in the group membership table, the device sends the join message or leave message of a program to the upper layer through the multicast uplink port. In addition, the device responds to the queries of the multicast router according to the status of the group membership table.

IGMP proxy effectively reduces the quantity of IGMP messages exchanged on the network side and therefore lessens the load of multicast routers. It is configurable on the device whether to send the IGMP general query to all multicast users or to only interested multicast users.

NOTE

After a multicast user orders a program for the first time, the system generates a value, indicating the random interval. The value ranges from 0 to the difference between the interval for the general query and maximum response time to the general query. After the random interval, the system sends the first general query message to the multicast user.

IGMP Snooping

IGMP snooping enables the device to listen to the multicast protocol packets transmitted between the router and hosts and set up a Layer 2 forwarding table for multicast data packets. This manages and controls multicast data packet forwarding so that the multicast data packets can be sent only to specified recipients. IGMP snooping features:

- Low bandwidth requirements, facilitating separate host charging.
- Separate data forwarding for each VLAN, improving data security.

IGMP snooping has two types:

- IGMP transparent snooping

  It is a snooping function without proxy. The device selects the proxy, snooping, or snooping with proxy function based on MVLANs.

  Before enabling IGMP transparent snooping, run the **igmp query-proxy**, **igmp report-proxy**, and **igmp leave-proxy** commands to disable proxy for query, report, and leave packets, respectively.

  IGMP transparent snooping enables the device to process IGMP messages as follows:

  – Query message

  After receiving the general query message and group-specific query message from the multicast uplink port, the device forwards the network-side query message to the user.

  NOTE
  - To ensure that the multicast user responds to the query in a timely manner, the maximum response time configured on the device must be shorter than that configured on the upper-layer multicast router.
  - The network-side IGMP version of the device is not affected by the multicast router.

  – Join/Leave message

  The device transparently transmits all the join/leave message received from the multicast user to the MVLAN.

  NOTE
  The IGMPv3 message may contain multiple group records that match different MVLANs. In this case, the device segments the message and transparent transmits the segmented messages to the corresponding MVLANs.

- IGMP snooping with proxy

  In IGMP upstream transmission, IGMP snooping with proxy is the same as IGMP proxy; in IGMP downstream transmission, however, IGMP snooping with proxy does not suppress the query message as IGMP proxy does.

  IGMP snooping with proxy enables the device to process IGMP messages as follows:

  – Query message

  After receiving the query message from the multicast uplink port, the device reconstructs and sends the query message to the user (default mode, query proxy enabled based on MVLANs) or forwards the query message (disabling query proxy based on MVLANs). It also responds to the multicast router's query according to its multicast group membership table.

  NOTE
  - Like IGMP proxy, the network-side IGMP version of the device is affected by the multicast router.
  - Determine whether to enable query proxy. For details, see the "Usage Guidelines" of the **igmp query-proxy** command.

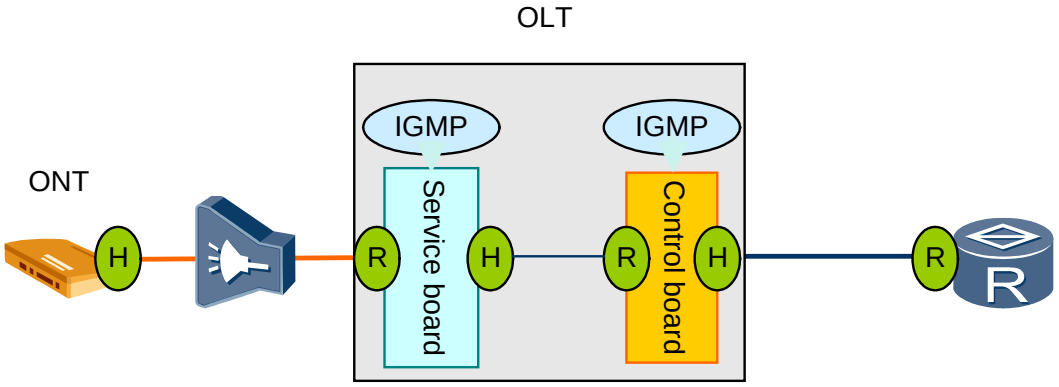
- Join/Leave message

The device sends only the first join message from the multicast users to the MVLAN. The device sends only the last leave message from the multicast users to the MVLAN.

IGMP Framework on the Device

Distributed IGMP [OLT]

Figure 18-13 Distributed protocol model



R represents the router functions of the IGMP protocol, and H represents the host functions of the IGMP protocol.

In the distributed two-level IGMP protocol stack, the first level is on the control board and the operation on the user side and the network side is based on MVLAN; the second level is on the service board, the operation on the network side is based on MVLAN, and the operation on the user side is based on multicast user, which ensures that users do not affect each other on the control plane. The convergence of the IGMP protocol stack on the service board lightens the processing load of the IGMP protocol stack on the control board. Given the same hardware conditions, the system can process channel switching of more multicast users at the same time.

NOTE

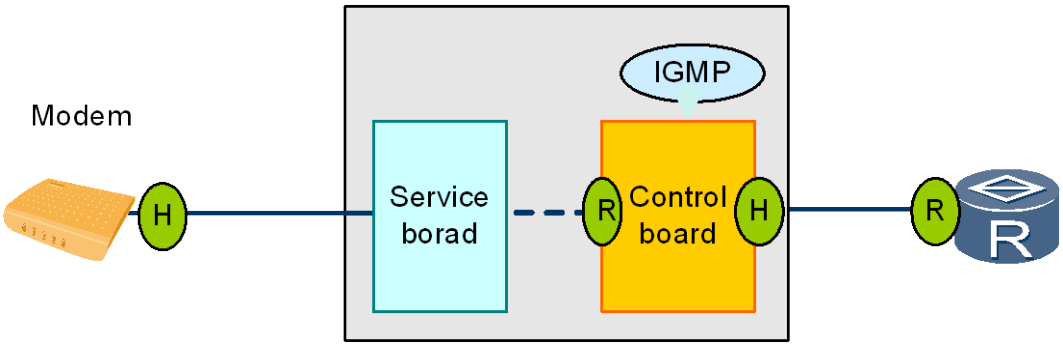
In the distributed multicast mode, the multicast configuration command and multicast query command are processed through different channels so that they can be executed concurrently. If a configuration command takes a long time to be executed and the corresponding query command is executed subsequently, the result of the query command may not display configuration result, depending on when the query command is executed.

For example, if you run the **igmp statistic reset vlan 100** command to clear the IGMP packet statistics and then run the **display igmp statistic vlan 100** command to query the IGMP packet statistics of multicast VLAN 100:

- When the **igmp statistic reset vlan 100** command execution is not completed yet and IGMP packet statistics are not cleared, the query result is not displayed as zero.
- When the **igmp statistic reset vlan 100** command is executed and IGMP packet statistics are cleared, the query result is displayed as zero.

Architecture of the IGMP Protocol Stack [DSLAM]

Figure 18-14 Architecture of the IGMP protocol stack



R represents the router functions of the IGMP protocol, and H represents the host functions of the IGMP protocol.

In the architecture of the IGMP protocol stack, the operation on the control board and the network side is based on MVLAN, and the operation on the user side is based on multicast user (to ensure that the control planes of the users do not interfere with each other).

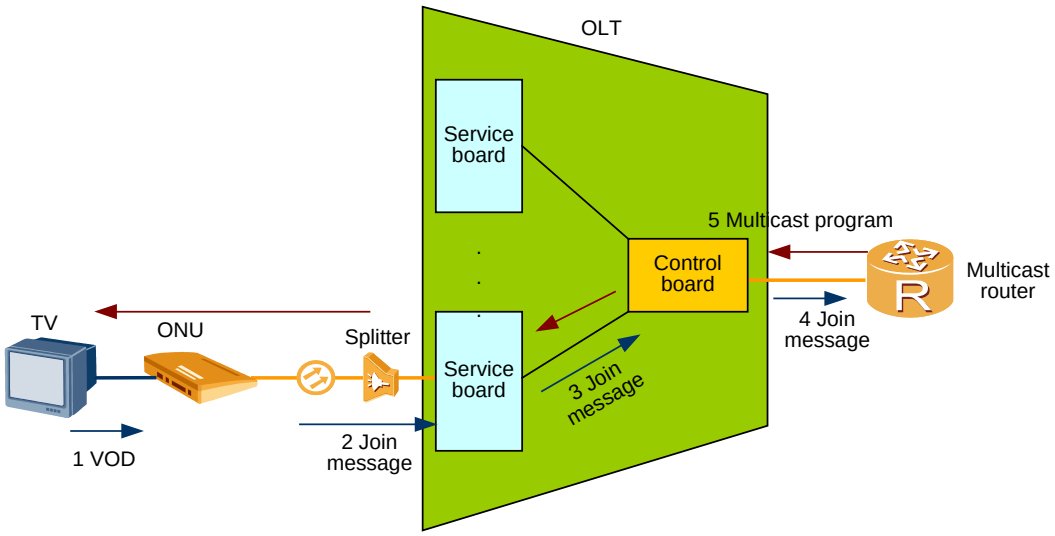
18.4.2 Mutlicast Forwarding

Join Flow

This section considers IGMPv2 proxy as an example to describe the Join Flow.

Join Flow [OLT]

Figure 18-15 Join flow



1. The multicast user switches a channel and sends a join message for demanding a new program GIP1.
2. After receiving the join message, the service board enters the IGMP protocol stack of the multicast user. After multicast control is implemented (for details, see "Multicast CAC"), the following group membership table is generated on the service board.

Index	Online Member
MVLAN1+GIP1	Multicast user 1

   - At the same time, the following multicast forwarding table is generated on the service board (for details on how to map GIP1 to GMAC1, see "Basic Concepts").

Index	Duplication Destination
MVLAN1+ GMAC1	GPON port 1

   - According to MVLAN1 corresponding to the program, the service board serves as the proxy of multicast user 1 and sends a join message to the control board.
3. After receiving the join message, the control board enters the IGMP protocol stack of MVLAN1 and generates the following group membership table.

Index	Online Member
MVLAN1+GIP1	Service board 1

   - At the same time, the control board generates the following multicast forwarding table.

Index	Duplication Destination
MVLAN1+ GMAC1	Port corresponding to service board 1

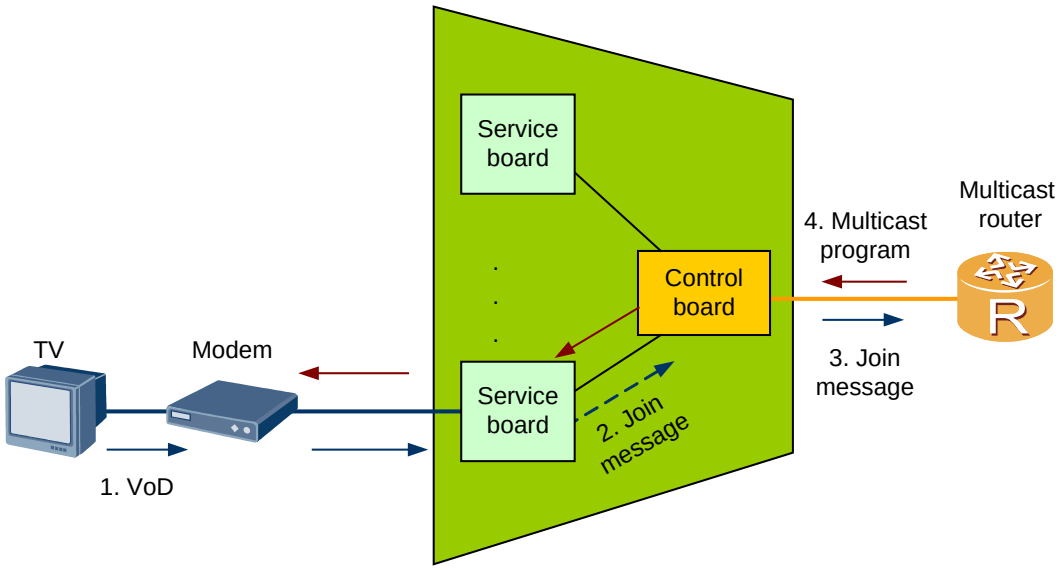
4. The control board then sends a join message to the multicast router through the multicast upstream port of MVLAN1.
5. After receiving the multicast stream, the device first duplicates the stream to service board 1 according to the multicast forwarding table of the control board, and then duplicates the stream to GPON port 1 according to the multicast forwarding table of the service board.

NOTE

1. Though the SVLAN of a multicast user is different from the MVLAN, the device can still implement the mapping to the MVLAN according to the multicast member configuration relationship. In this way, cross-VLAN multicast is supported without requiring additional configuration.
2. The join flow for boards supporting the group filter mode is similar and only the forwarding entry index is different.

Join Flow [DSLAM]

Figure 18-16 Join flow



1. The multicast user switches a channel and sends a join message for demanding a new program GIP1.
2. After receiving the join message, the control board enters the IGMP protocol stack of the multicast user. After multicast control is implemented (for details, see "Multicast CAC"), the following group membership table is generated on the control board.

Index	Online Member
MVLAN1+GIP1	Multicast user 1

At the same time, the following multicast forwarding tables are generated on the control board and the service board (for details on how to map GIP1 to GMAC1, see "Basic Concepts").

Table 18-6 Multicast forwarding table on the control board

Index	Duplication Destination
MVLAN1+ GMAC1	Corresponding port on service board 1

Table 18-7 Multicast forwarding table on the service board

Index	Duplication Destination
MVLAN1+ GMAC1	User port 1

- According to MVLAN1 corresponding to the program, the service board serves as the proxy of multicast user 1 and sends a join message to the control board.

3. The control board then sends a join message to the multicast router through the multicast upstream port of MVLAN1.
4. After receiving the multicast stream, the device first duplicates the stream to service board 1 according to the multicast forwarding table of the control board, and then duplicates the stream to user port 1 according to the multicast forwarding table of the service board.

NOTE

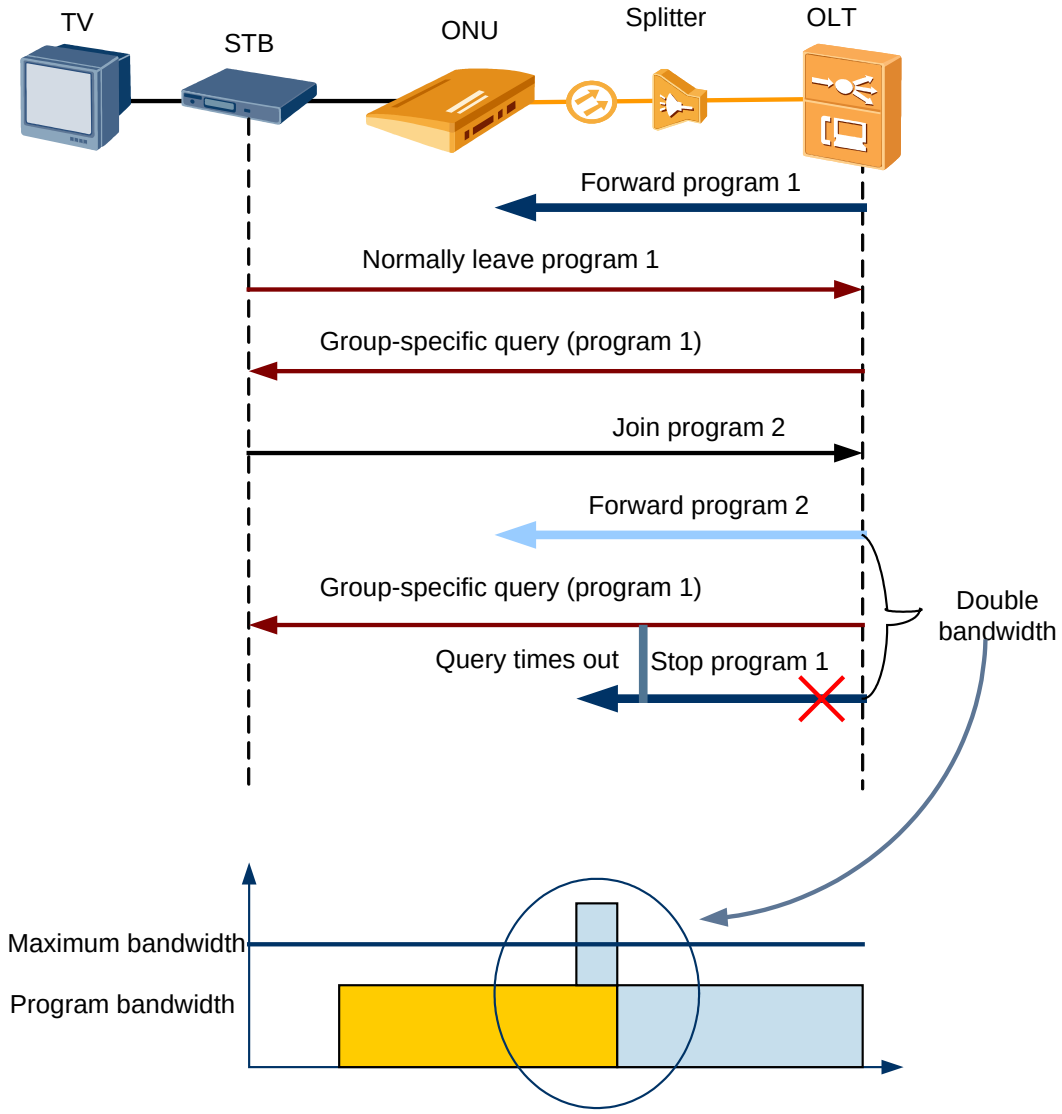
1. Though the SVLAN of a multicast user is different from the MVLAN, the device can still implement the mapping to the MVLAN according to the multicast member configuration relationship. In this way, cross-VLAN multicast is supported without requiring additional configuration.
2. The join flow for boards supporting the group filter mode is similar and only the forwarding entry index is different.
3. If the traffic with a high priority is suddenly overloaded and the service with a low priority is affected, IGMP packets are not discarded. The device processes and sends the IGMP packets first.

Leave Flow

Normal leave

As defined by IGMPv2, the router must send a group-specific query message after it receives the leave message from a host, and it considers that the host does not need the data of the group until the query times out. The following figure illustrates the flow of a normal leave (the same to IGMPv3).

Figure 18-17 Flow of a normal leave

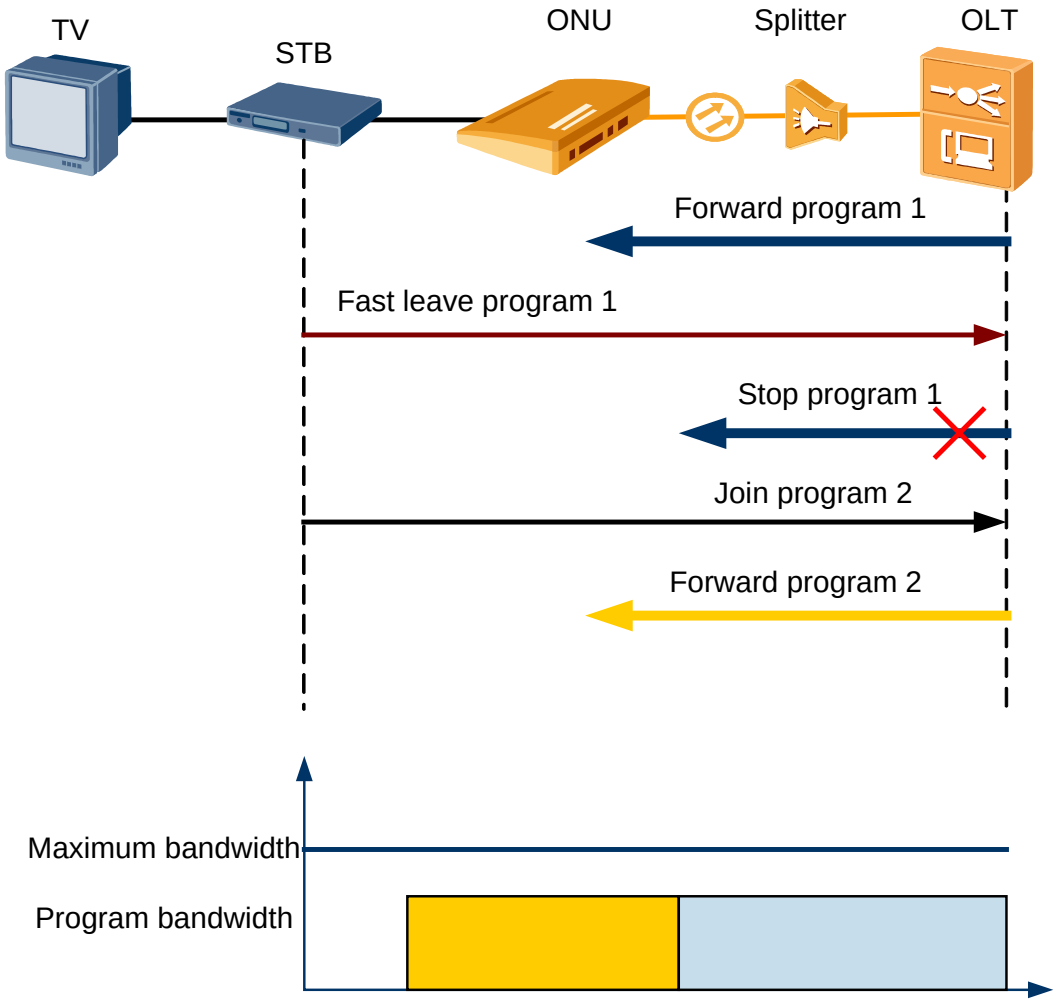


Two IGMP messages are transmitted in the case of a channel switching, one for leaving the original multicast group and one for joining the new multicast group. Therefore, traffic of two multicast groups exists on the subscriber line before the original multicast group is stopped. If the subscriber line does not reserve sufficient bandwidth for carrying the traffic of two multicast groups, traffic overflow (packet loss) will occur. For example, if video streams are carried, pixelation will occur.

Fast leave

When the device receives the leave message from a multicast user, it immediately stops forwarding the messages of the user. The following figure illustrates the flow of a fast leave (the same to IGMPv3).

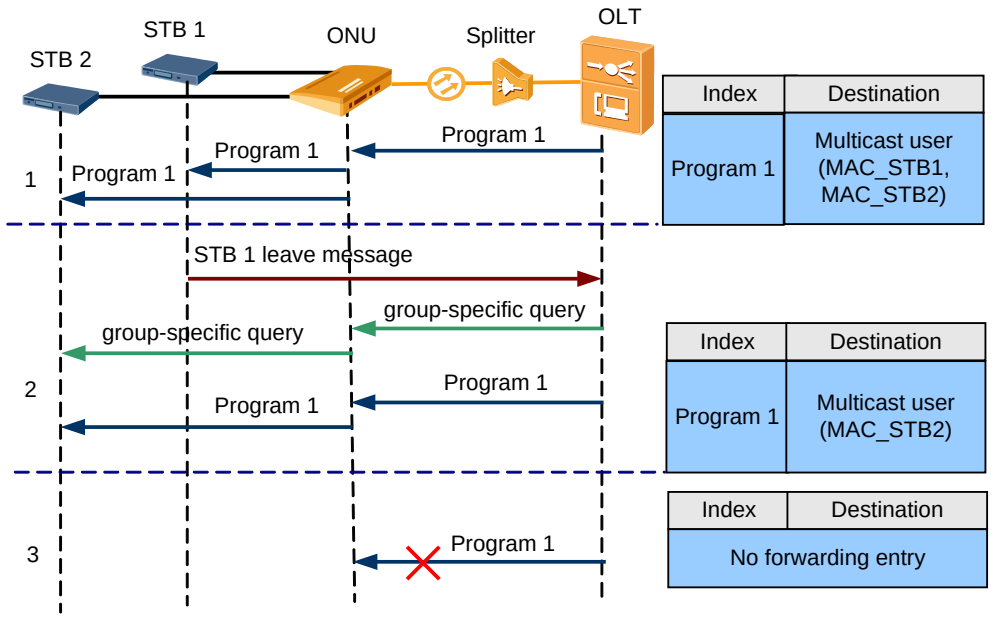
Figure 18-18 Flow of a fast leave



Fast leave based on the MAC address

When the device generates a multicast group membership table, it not only records the multicast user but also records the MAC addresses of the multicast group members of the multicast user. A maximum of eight MAC addresses are supported for each multicast user. When the device receives a leave message, it first deletes the MAC addresses in the multicast group membership table, and it stops forwarding the messages of the group only when all the MAC addresses of the multicast user are deleted. The following figure illustrates the flow of a fast leave based on the MAC address (the same to IGMPv3).

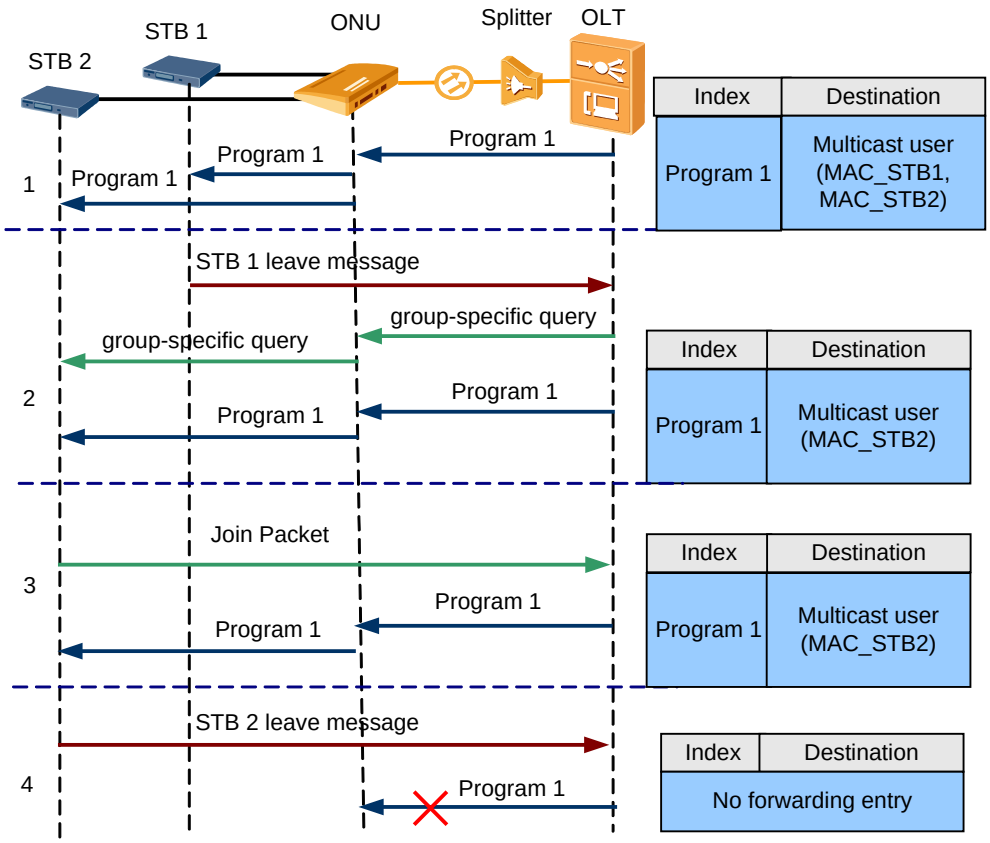
Figure 18-19 Flow of a fast leave based on the MAC address (STB2 has silently left.)



NOTE

The silent leaving of an STB indicates that the STB exceptionally leaves a program. For example, the STB is powered off or goes offline.

Figure 18-20 Flow of a fast leave based on the MAC address (STB2 is online.)



In summary, the three leave modes have their advantages and disadvantages. You can choose any of them according to actual needs and can set the leave mode for a multicast user.

	User-side Multi-STB Supported or Not	Bandwidth Occupation Time
Normal leave	Yes The STB quantity is not limited.	Aged upon reception of group-specific query
Fast leave	No	Released immediately
Fast leave based on the MAC address	Yes One multicast user supports a maximum of eight STBs at a time.	Released immediately

The following configurations are recommended and can be adopted by different users according to their home network topologies.

HG Function	STB Quantity	Reserved Bandwidth	Normal Leave	Fast Leave	Fast Leave Based on the MAC Address

HG Function	STB Quantity	Reserved Bandwidth	Normal Leave	Fast Leave	Fast Leave Based on the MAC Address
No IGMP	One	Insufficient		√	√
		Sufficient	√	√	√
	Several	Insufficient			√ (Less than eight)
		Sufficient	√		√ (Less than eight)
IGMP snooping	One	Insufficient		√	√
		Sufficient	√	√	√
	Several	Insufficient			√ (Less than eight)
		Sufficient	√		√ (Less than eight)
IGMP proxy	One	Insufficient		√	√ (Not limited)
		Sufficient	√	√	√ (Not limited)
	Several	Insufficient		√	√ (Not limited)
		Sufficient	√	√	√ (Not limited)

Global Leave

As defined in TR101, the global leave message is an IGMP message with an all-zero group IP address, which indicates leaving all the groups.

- Network side

When the network topology changes, the device sends the global leave message to the upper-layer multicast router. After receiving the message, the upper-layer multicast router immediately sends the general query message, with the maximum response time set to the maximum time of responding to the group-specific query message. The device, after receiving the query message, responds to the upper-layer multicast router with the join message of the interested group. In this way, the multicast service can recover more quickly. Here, the network topology change events include ring network switching, line up/down, and active/standby port switching in a protect group.

NOTE

- If the device is interconnected with a network device that does not support the global leave message, multicast services may be interrupted during the network topology change. Therefore, it is recommended that the global leave function be manually disabled on the device.
- The device supports sending of the global leave message only in IGMPv2.

- User side

When the STB is powered on immediately after a sudden power-off, because the STB cannot remember the previously-watched program, the bandwidth of the previously-watched program and the program resources are released only after the general query ages.

If the STB supports the global leave function, the STB sends a global leave message after it is re-powered on. After receiving the message, the device sends a general query message, with

the maximum response time set to the maximum time of responding to the group-specific query message. If the multicast user is a fast-leave or MAC-based fast-leave user, the device releases all program resources of this multicast user. If the user is a normal-leave user, the device sends a group-specific query message and releases the program resources after the group-specific query times out.

NOTE
Only the IGMPv2 global leave messages can be processed.

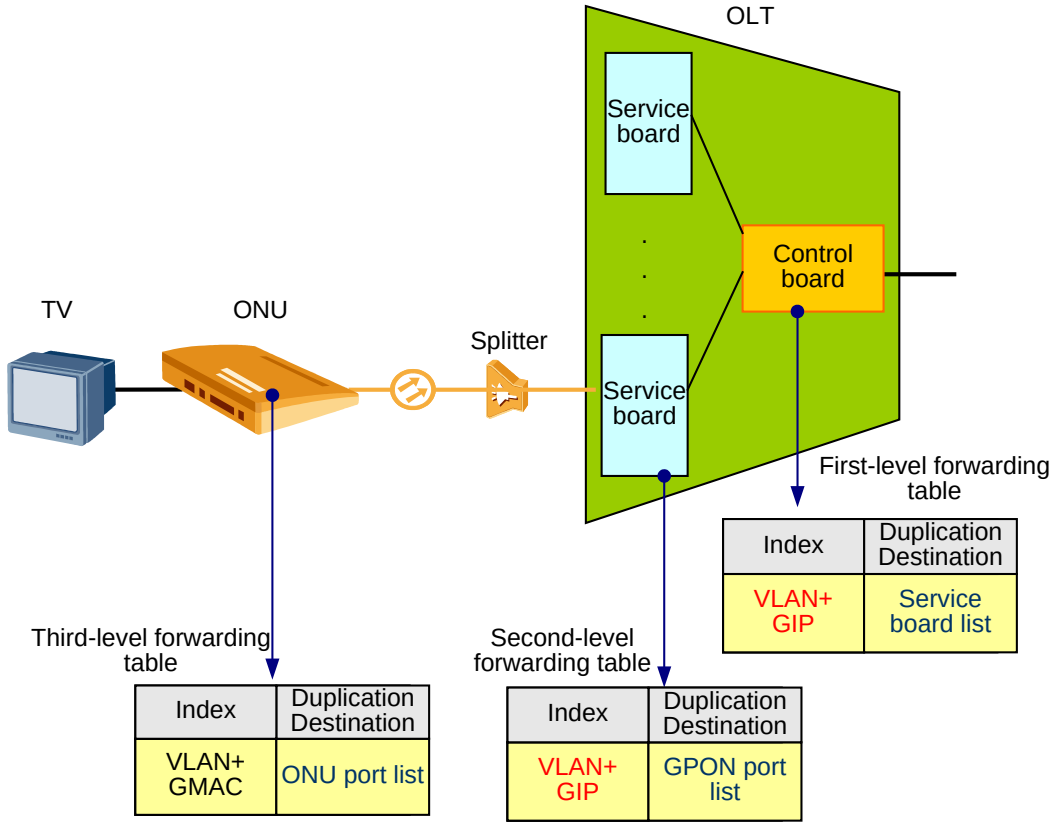
Forwarding Framework on the Device

For OLT device

NOTE
The difference among the following forwarding tables is the first-level and second-level indexes. For example, the first-level and second-level index in the GPON multicast forwarding table (VLAN+GIP) is VLAN+GIP.

GPON Multicast Forwarding Table (VLAN+GIP) [OLT]

Figure 18-21 GPON multicast forwarding table (VLAN+GIP)



The OLT supports a distributed 2+1-level duplication architecture:

- The first-level duplication is implemented on the control board. By using the "VLAN+GIP" index, the control board duplicates multicast data to the service board interested in the multicast program in an as-per-requirement manner, effectively saving the backplane bandwidth.

- The second-level duplication is implemented on the service board. By using the "VLAN+GIP" index, the service board duplicates multicast data to the GPON port interested in the multicast program in an as-per-requirement manner, effectively saving the downstream bandwidth of the GPON port. Then the service board encapsulates and transmits the multicast data on the GPON port in the mode of multicast GEM port (system-level parameter, configurable, default value 4095).
- The third-level duplication is implemented on the ONT. By using the "VLAN+GMAC" white list, the ONT filters out unneeded multicast data to avoid bandwidth overflow at the downstream ingress (ONT only supports that in olt-control mode). Then, by using the "VLAN+GMAC" index, the ONT duplicates the multicast data to the ONT ports in an as-per-requirement manner (only supports forwarding by using GMAC in snooping mode).

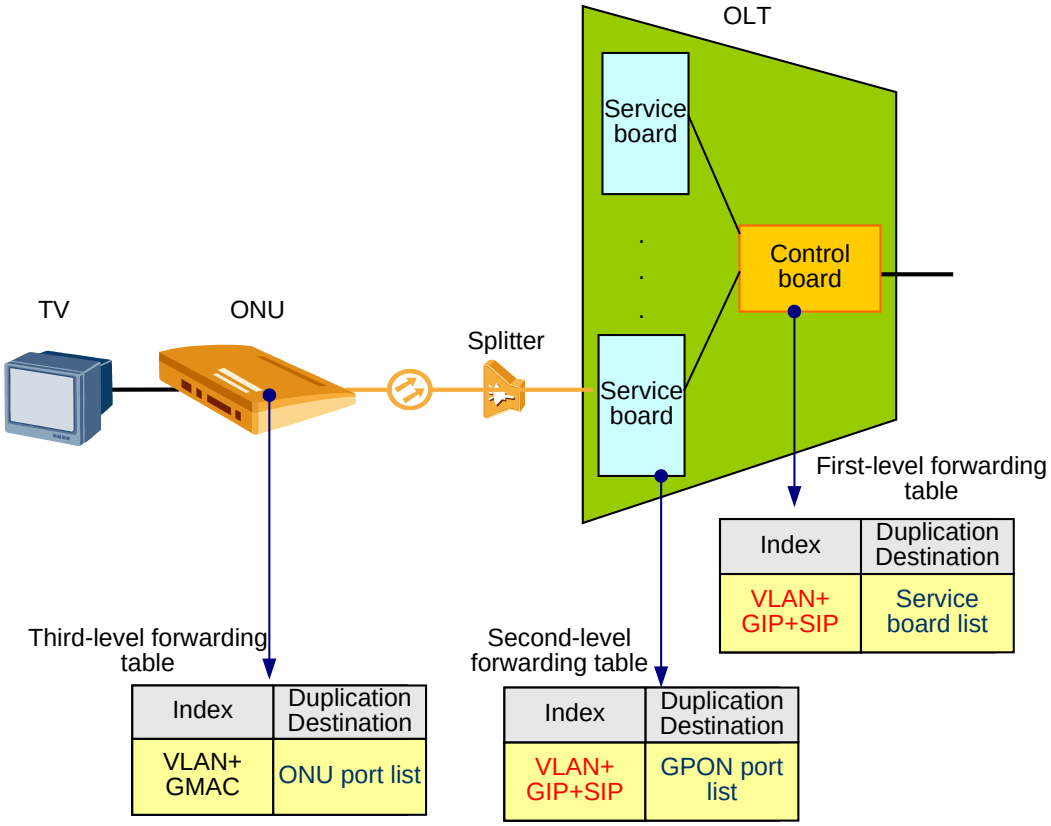
NOTE

1. This forwarding mode applies to ASMSSM and ASM ONLY group filter modes (for details, see "18.3 Multicast Model").
2. The ONT that implements the third-level duplication is recommended to use the chip that supports VLAN+GIP forwarding. If the ONT does not use such a chip, program GIP-to-GMAC mapping must be a one-to-one mapping. Without the one-to-one mapping on the same ONT, garbled images will occur.

This topic describes only the forwarding framework in the most common single-copy duplication mechanism. For the hardware forwarding framework in the multi-copy duplication mechanism, see "GPON Multi-Copy Duplication."

GPON Multicast Forwarding Table (VLAN+GIP+SIP) [OLT]

Figure 18-22 GPON multicast forwarding table (VLAN+GIP+SIP)



The OLT supports a distributed 2+1-level duplication architecture:

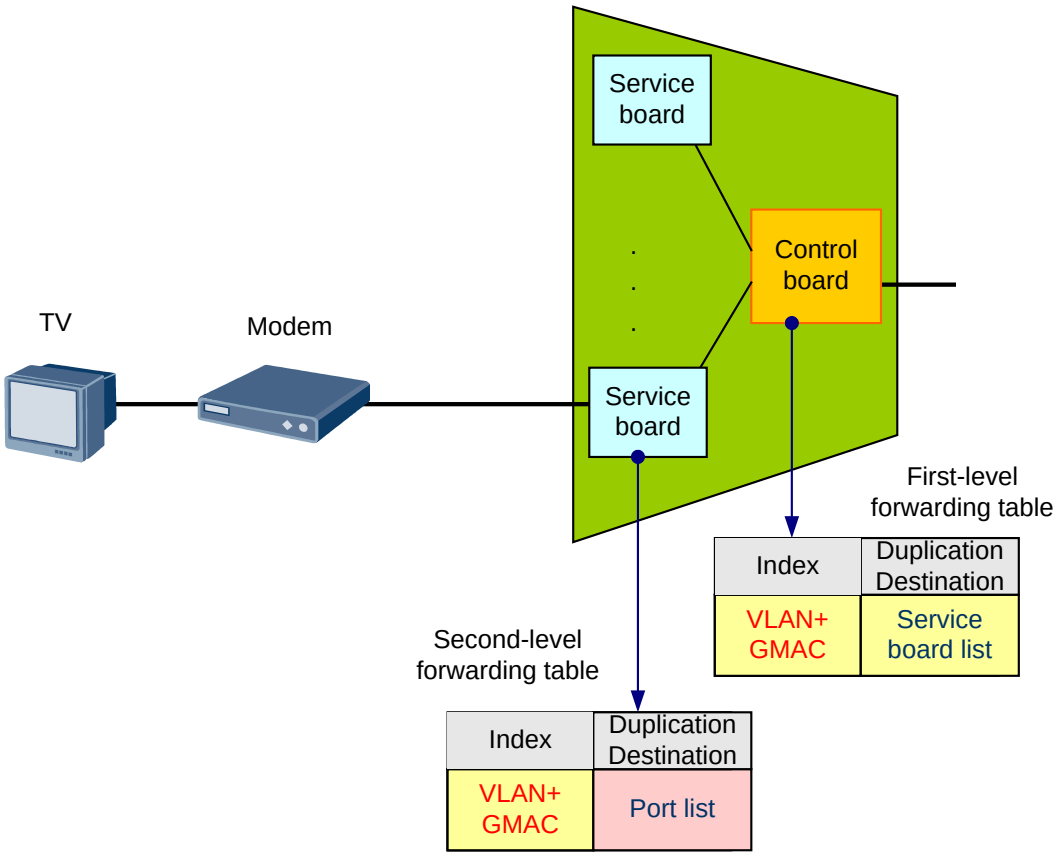
- The first-level duplication is implemented on the control board. By using the "VLAN+GIP+SIP" index, the control board duplicates multicast data to the service board interested in the multicast program in an as-per-requirement manner, effectively saving the backplane bandwidth.
- The second-level duplication is implemented on the service board. By using the "VLAN+GIP+SIP" index, the service board duplicates multicast data to the GPON port interested in the multicast program in an as-per-requirement manner, effectively saving the downstream bandwidth of the GPON port. Then the service board encapsulates and transmits the multicast data on the GPON port in the mode of multicast GEM port (system-level parameter, configurable, default value 4095).
- The third-level duplication is implemented on the ONT. By using the "VLAN+GMAC" white list, the ONT filters out unneeded multicast data to avoid bandwidth overflow at the downstream ingress (ONT only supports that in olt-control mode). Then, by using the "VLAN+GMAC" index, the ONT duplicates the multicast data to the ONT ports in an as-per-requirement manner (only supports forwarding by using GMAC in snooping mode).

NOTE

1. This forwarding mode applies to the SSM ONLY group filter mode.
2. Due to hardware limitations, only the difference in the least significant 20 bits can be differentiated between source IP addresses (SIPs). For example, 1.1.1.1 and 2.1.1.1 are the same SIP for the device; 1.1.1.1 and 1.1.1.2 are different SIPs for the device.
3. The ONT that implements the third-level duplication is recommended to use the chip that supports VLAN+GIP forwarding. If the ONT does not use such a chip, program GIP-to-GMAC mapping must be a one-to-one mapping. Without the one-to-one mapping on the same ONT, garbled images will occur.

Multicast Forwarding Table (VLAN+GMAC) [DSLAM]

Figure 18-23 DSLAM multicast forwarding table (VLAN+GMAC)



The DSLAM supports a distributed two-level duplication architecture:

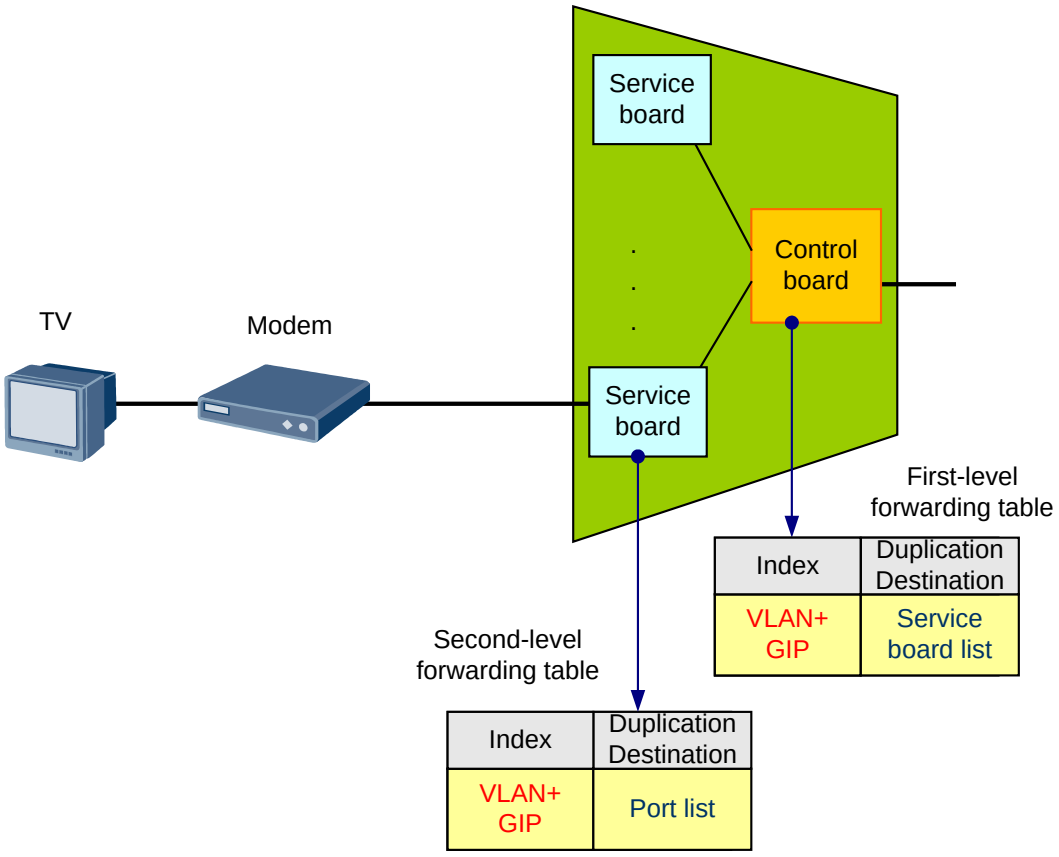
- The first-level duplication is implemented on the control board. By using the "VLAN+GMAC" index, the control board duplicates multicast data to the service board interested in the multicast program in an as-per-requirement manner, effectively saving the backplane bandwidth.
- The second-level duplication is implemented on the service board. By using the "VLAN+GMAC" index, the service board duplicates multicast data to the multicast user (usually corresponding to the first port) interested in the multicast program in an as-per-requirement manner.

NOTE

This forwarding mode applies to boards that do not support Layer 3 multicast forwarding chip.

Multicast Forwarding Table (VLAN+GIP) [DSLAM]

Figure 18-24 DSLAM multicast forwarding table (VLAN+GIP)



The DSLAM supports a distributed two-level duplication architecture:

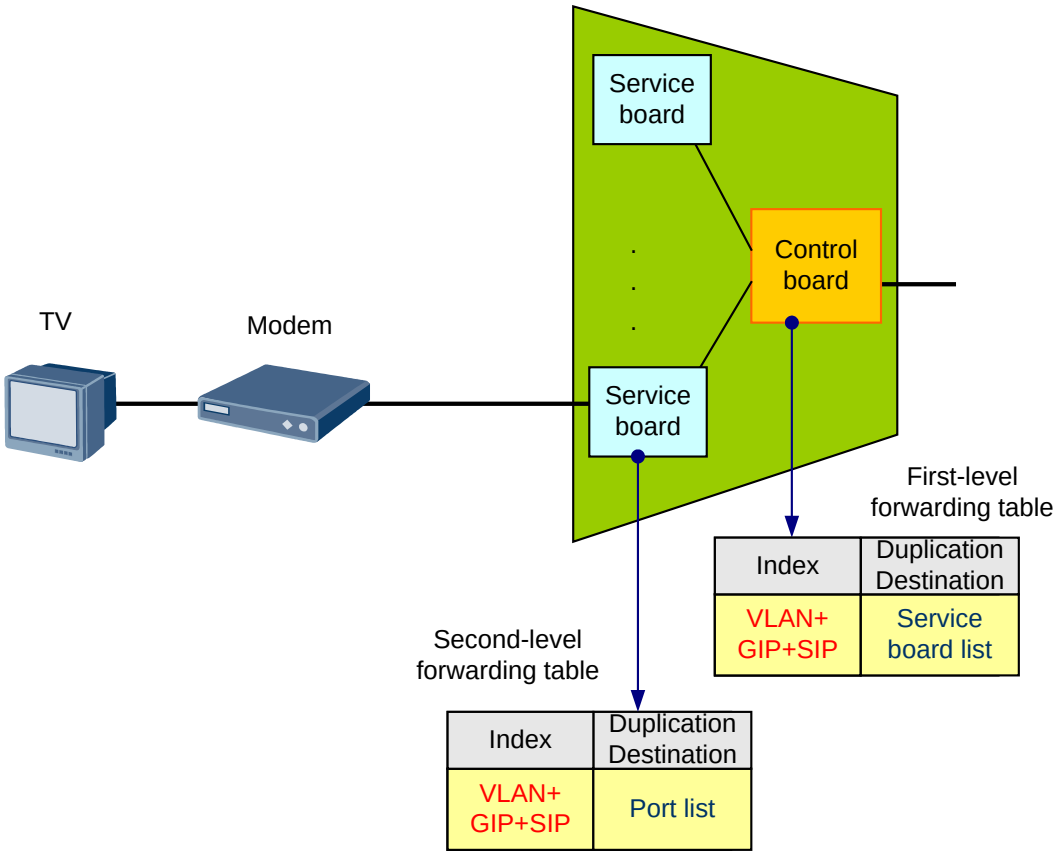
- The first-level duplication is implemented on the control board. By using the "VLAN+GIP" index, the control board duplicates multicast data to the service board interested in the multicast program in an as-per-requirement manner, effectively saving the backplane bandwidth.
- The second-level duplication is implemented on the service board. By using the "VLAN+GIP" index, the service board duplicates multicast data to the multicast user (usually corresponding to the first port) interested in the multicast program in an as-per-requirement manner.

NOTE

This forwarding mode applies to ASMSSM and ASM ONLY group filter modes.

Multicast Forwarding Table (VLAN+GIP+SIP) [DSLAM]

Figure 18-25 DSLAM multicast forwarding table (VLAN+GIP+SIP)



The DSLAM supports a distributed two-level duplication architecture:

- The first-level duplication is implemented on the control board. By using the "VLAN+GIP+SIP" index, the control board duplicates multicast data to the service board interested in the multicast program in an as-per-requirement manner, effectively saving the backplane bandwidth.
- The second-level duplication is implemented on the service board. By using the "VLAN+GIP+SIP" index, the service board duplicates multicast data to the multicast user (usually corresponding to the first port) interested in the multicast program in an as-per-requirement manner.

NOTE

1. This forwarding mode applies to the SSM ONLY group filter mode.
2. Due to hardware limitations, only the difference in the least significant 20 bits can be differentiated between SIPs. For example, 1.1.1.1 and 2.1.1.1 are the same SIP for the device; 1.1.1.1 and 1.1.1.2 are different SIPs for the device.

GPON Multicast Duplication

If the service board of the device is a GPON service board, the device has two multicast forwarding mechanisms and you can configure the forwarding mechanism based on MVLAN.

- Single-copy duplication

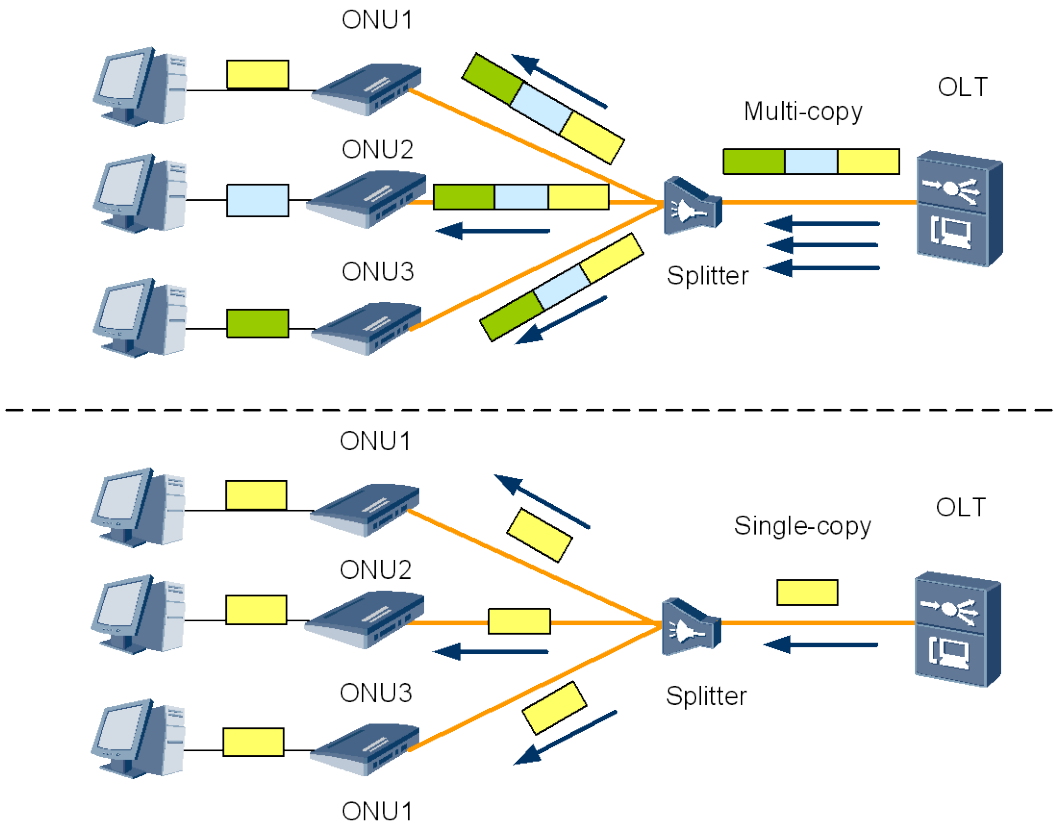
This is the most common duplication mode of GPON multicast (it also refers to the mode mentioned in this document unless otherwise stated). This mode makes the best of the GPON downstream WDM transmission principle and sends multicast data to all ONTs using the non-encrypted GEM port bearer channel. Each ONT receives the multicast data according to the multicast filtering table. For the corresponding hardware forwarding entry, see "Forwarding Framework on the Device."

- Multi-copy duplication

  The primary difference between single-copy duplication and multi-copy duplication is that, in multi-copy duplication, multicast data is duplicated to corresponding service ports according to user requirements, encapsulated in the encrypted unicast GEM port channel, and then sent to the ONT. The following table shows the multicast forwarding table at the GPON board level.

Index	Entry
VLAN+GMAC	Multicast user list

Figure 18-26 Single-copy duplication and multi-copy duplication



The following table lists the differences between single-copy duplication and multi-copy duplication.

	Single-copy Duplication	Multi-copy Duplication

	Single-copy Duplication	Multi-copy Duplication
Duplication granularity	Based on GPON port	Based on multicast user
Bandwidth	One GPON port has only one multicast stream.	One multicast user has only one multicast stream, but one GPON port may have multiple multicast streams.
Security	On the one hand, the security depends on the ONT filtering; on the other hand, the head end and STB encryption system are required.	This mode uses the GPON line AES128 encryption system and the real-time key conversion function, which provides better security than the common encryption system of the head end.
CAC of PON port	Supported	Not supported

18.4.3 Multicast Upstream Interoperation

Multicast Cascading

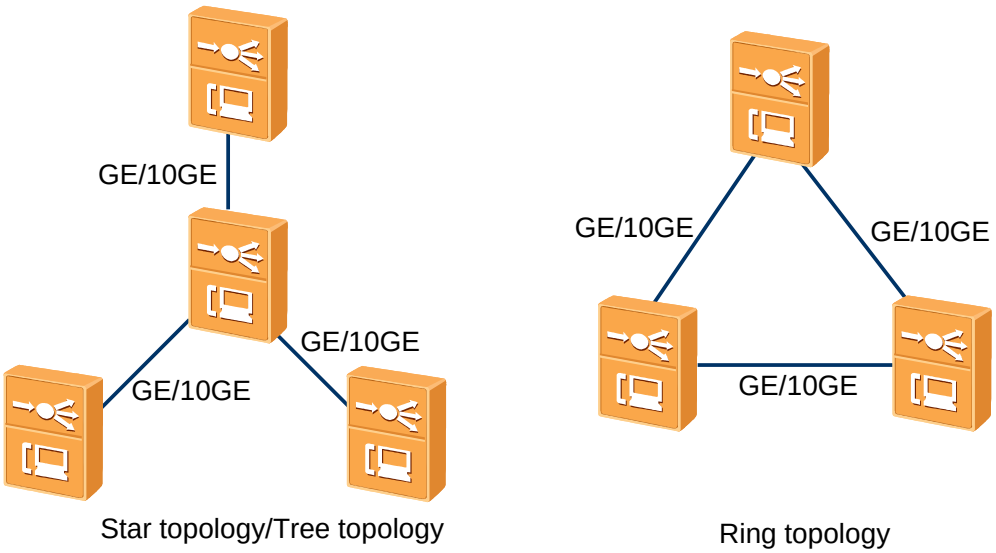
OLT Multicast Cascading

- Ethernet cascading

  Using Ethernet cascading on the access device, the number of ports on the convergence device and the optical cable routing cost can be reduced. In addition, capacity expansion for more users in the residential community access area can be easily implemented.

  There are two common cascading network topologies, star (tree/chain) network and ring network, as shown in the following figure. Here, the star cascading network is used as an example. For details about the ring cascading network, see "Ring Network of Uplink Ports".

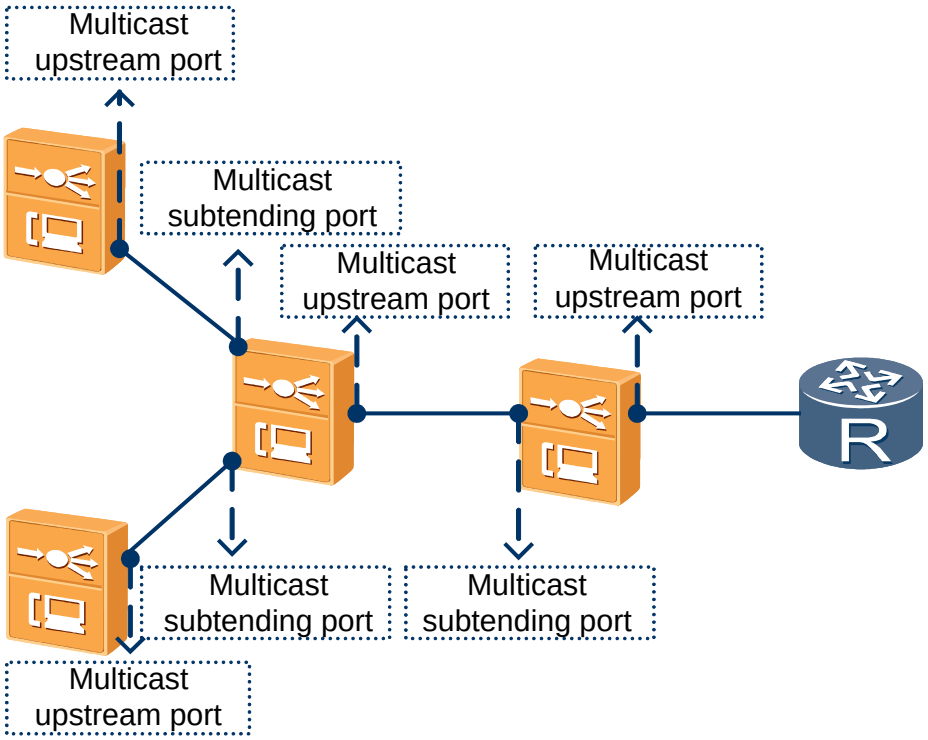
Figure 18-27 Cascading network topologies



- Multicast cascading port configuration

  The device, through the Ethernet port of its GIU board or other cascading boards, can be physically connected to the lower-layer device. Multicast service is configured through the multicast cascading ports, and in this way the interoperation between the devices is managed. A multicast cascading port corresponds to a physical port (the channel for carrying services can be created through the port VLAN or service stream). The following figure shows the relationship between the multicast cascading port and the multicast upstream port.

Figure 18-28 Multicast cascading port and upstream port

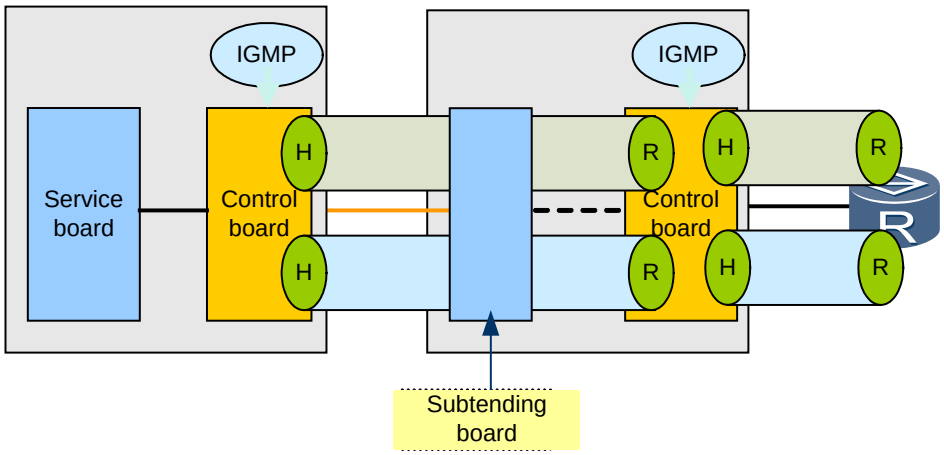


- IGMP control message

  In the multicast cascading scenario, the upper-layer device and the lower-layer device run the IGMP protocol stack separately. For a device, the cascading port (its lower-layer device) can be regarded as the multicast user. Multicast users are controlled by the lower-layer device and therefore the device does not support the following service functions for multicast users: rights management, multicast preview, multicast CAC, charging, and multicast service acceptance. The device supports the fast leave and normal leave functions.

  On the cascading port, the IGMP protocol stack is based on different VLANs, as shown in the following figure.

Figure 18-29 IGMP protocol stack on the cascading port



NOTE

If an Ethernet port is not configured as the multicast cascading port, the Ethernet port discards the IGMP report message.

The (SIP, GIP) field of the IGMP message and the VLAN of the IGMP message are used for program matching. The policy of processing unmatched messages can be configured based on the cascading port.

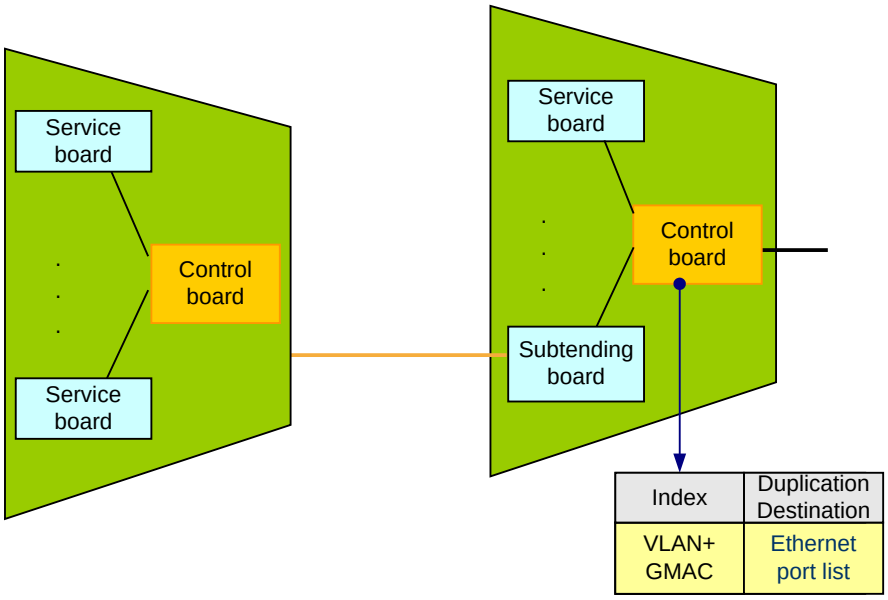
Considering the IGMP processing performance of the source node, it is recommended that all cascading devices adopt IGMP proxy instead of IGMP snooping.

- Multicast data forwarding

  Multicast data can be forwarded only in a VLAN. According to different cascading boards, there are two forwarding architectures.

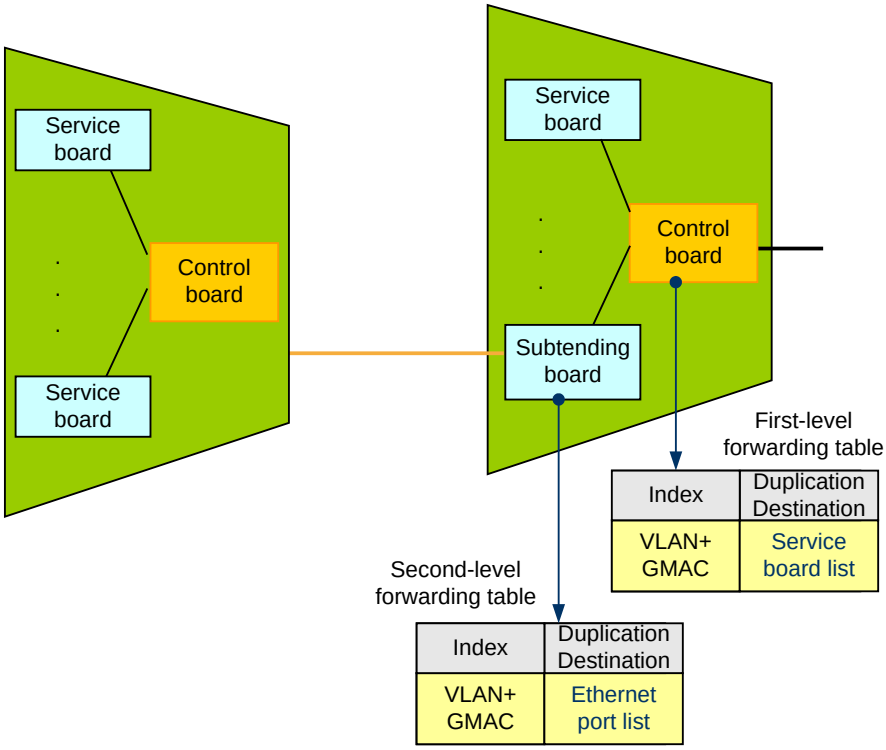
  - One-level forwarding architecture: The following figure uses the GIU board as an example.

Figure 18-30 One-level forwarding architecture of multicast cascading



- Two-level forwarding architecture: The following figure uses the ETHB/SPUA board as an example.

Figure 18-31 Two-level forwarding architecture of multicast cascading



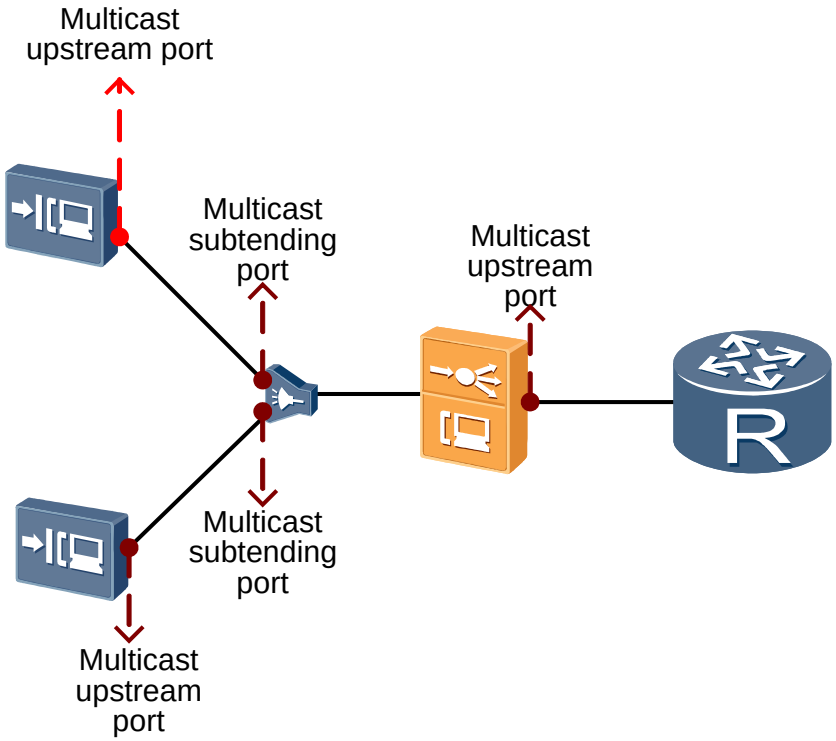
- xPON cascading [OLT]

  xPON cascading meets the requirements for multicast services in the FTTC/FTTB scenario.

  – Multicast cascading port configuration

    The device can implement the physical connection between the OLT and the MxU by using the PON line. Similar to Ethernet cascading mode, in xPON cascading mode, the interconnection between devices is also managed through the multicast cascading port object. One xPON cascading port corresponds to a logical interface (GEM port or LLID). The actual bearer channel can be created by using the service port. The following figure shows the relationship between the xPON cascading port and upstream port.

Figure 18-32 xPON cascading port and upstream port

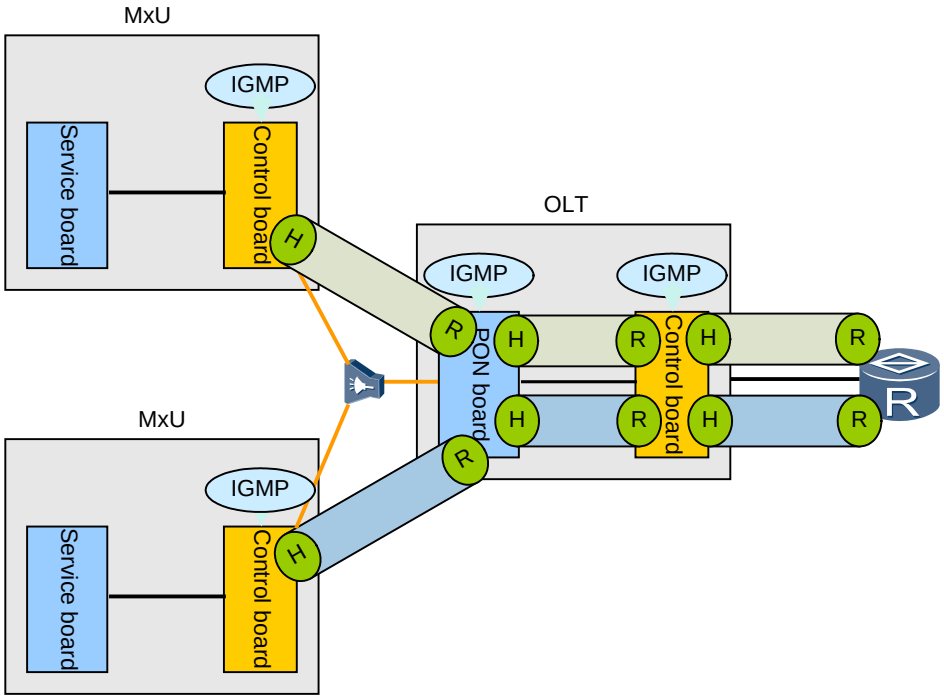


- IGMP control packet

  In the xPON cascading scenario, the OLT and the MxU run the IGMP protocol stack separately. Like Ethernet cascading, xPON cascading supports normal leave and fast leave.

  On the multicast cascading port, the IGMP protocol stack is based on different VLANs, as shown in the following figure. Different from Ethernet cascading, in xPON cascading, the bearer channel is limited by the maximum number of service ports that can be created on each GEM port or LLID, because the bearer channel is based on the service port. To support an MVLAN that is beyond the supported specifications, you can configure multiple GEM ports or LLIDs.

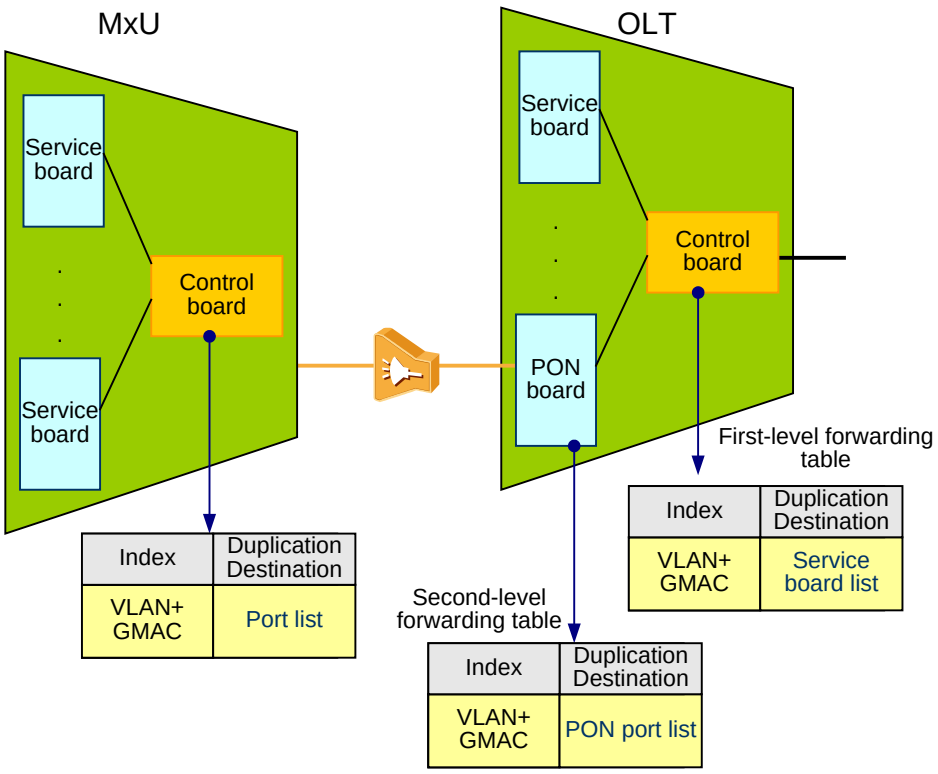
Figure 18-33 IGMP protocol stack of the xPON multicast cascading port



- Multicast data forwarding

  xPON cascading supports forwarding in the same VLAN and does not support cross-VLAN forwarding.

Figure 18-34 xPON multicast forwarding architecture



NOTE

Service ports that adopt traffic classification by two-tagged VLANs do not support multicast cascading ports.

Ring Network of Uplink Ports

In the ring network, access devices on the physical link are connected to form a ring. Devices on the ring maintain the ring status by running the Layer 2 link protocol.

The ring network of access devices has two advantages:

- Low network construction costs: In the ring network, an access device does not need to connect to the convergence switch, but connects to its nearest access device. This significantly saves optical cable resources. The switch provides only a few ports for the access device. In the ring network, however, deploying a small number of switches can meet the access requirements.
- High reliability: The Layer 2 link protocol provides the uplink backup protection. With this function, when the uplink of a single access device is faulty, the device can switch to the backup uplink.

The multicast service supports the following two ring networks on the network side.

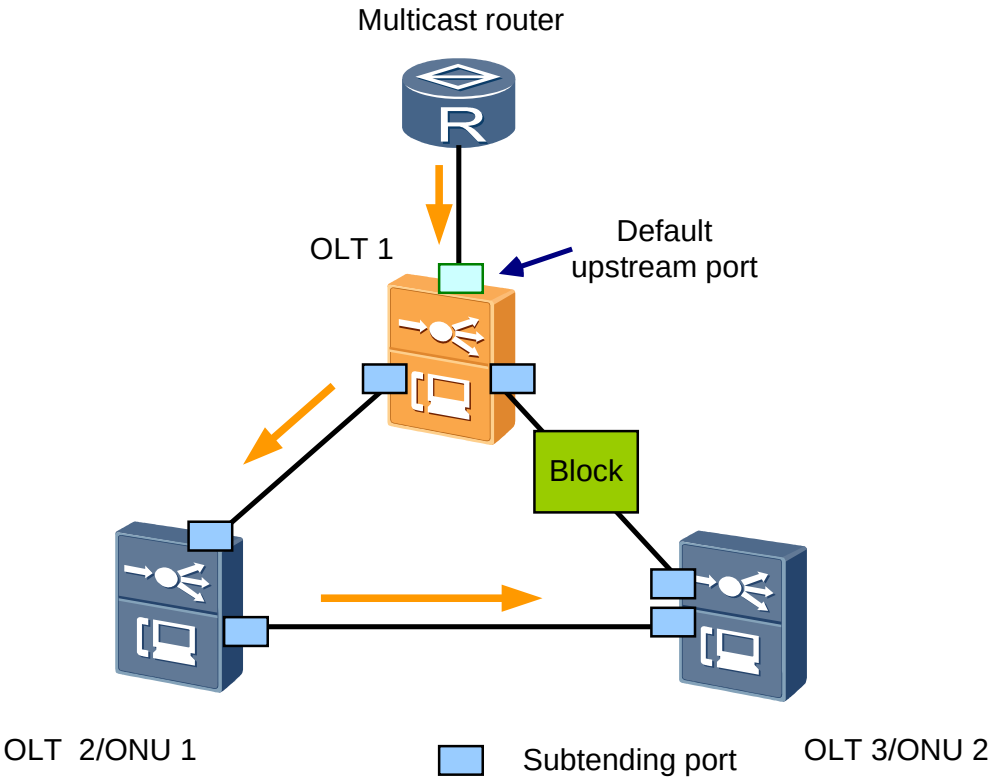
- MSTP

  The mode of the multicast uplink port needs to be set to MSTP by running the **igmp uplink-port-mode** command. In this case, the multicast uplink port of the device does not need to be configured; instead, the root port determined dynamically by MSTP serves as the multicast uplink port. If the access device is the MSTP root bridge (this device must be the injection point of the multicast data; using the MSTP priority

configuration, ensure that this root bridge is not removed), the access device does not have the root port. Therefore, in the actual network, the multicast data injection port needs to be configured as the default multicast uplink port.

In addition, the device ports on the ring need to be configured as the multicast cascading ports. The actual multicast downstream ports are determined by IGMP according to the multicast group membership table. The following figure shows the configuration of each role.

Figure 18-35 Multicast configuration in the MSTP ring network



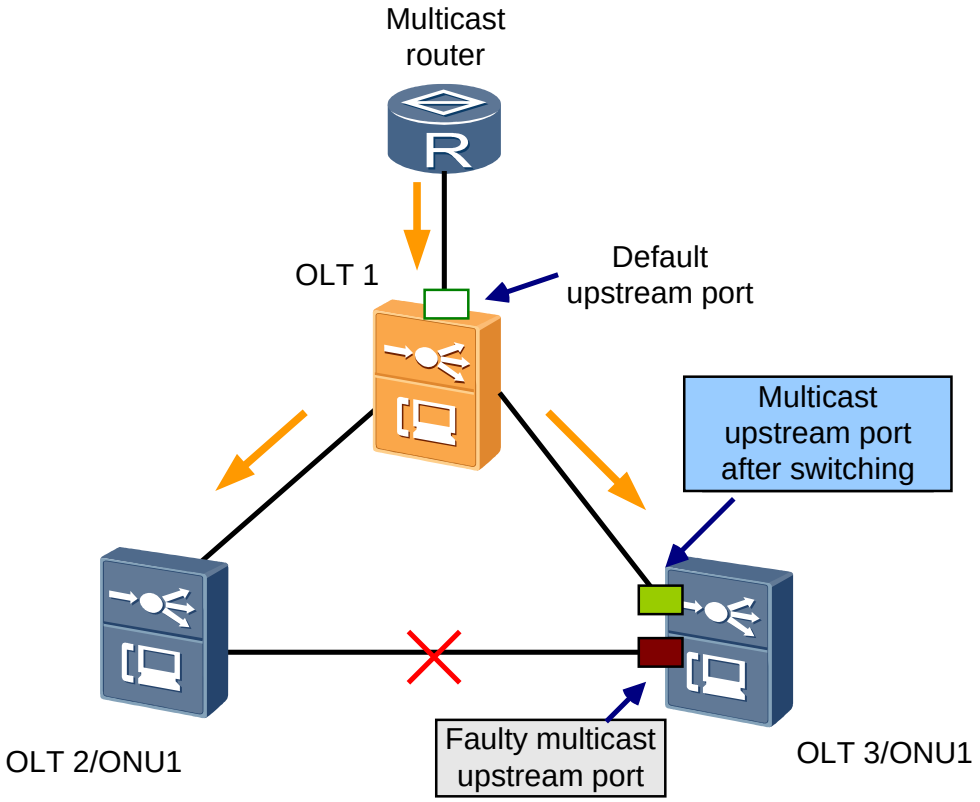
NOTE

The root bridge must be the injection point of the multicast service.

When the OLT using the ETHA/SPUA/ETHB board for upstream transmission, the device does not support MSTP multicast.

In the case of a link or device failure, after MSTP selects a backup link, the MVLAN-based IGMP protocol stack immediately sends the new root port (serving as the multicast uplink port) the join message targeting at the multicast group that the device is interested in. In this way, fast recovery of the multicast service can be ensured.

Figure 18-36 MSTP ring network fault

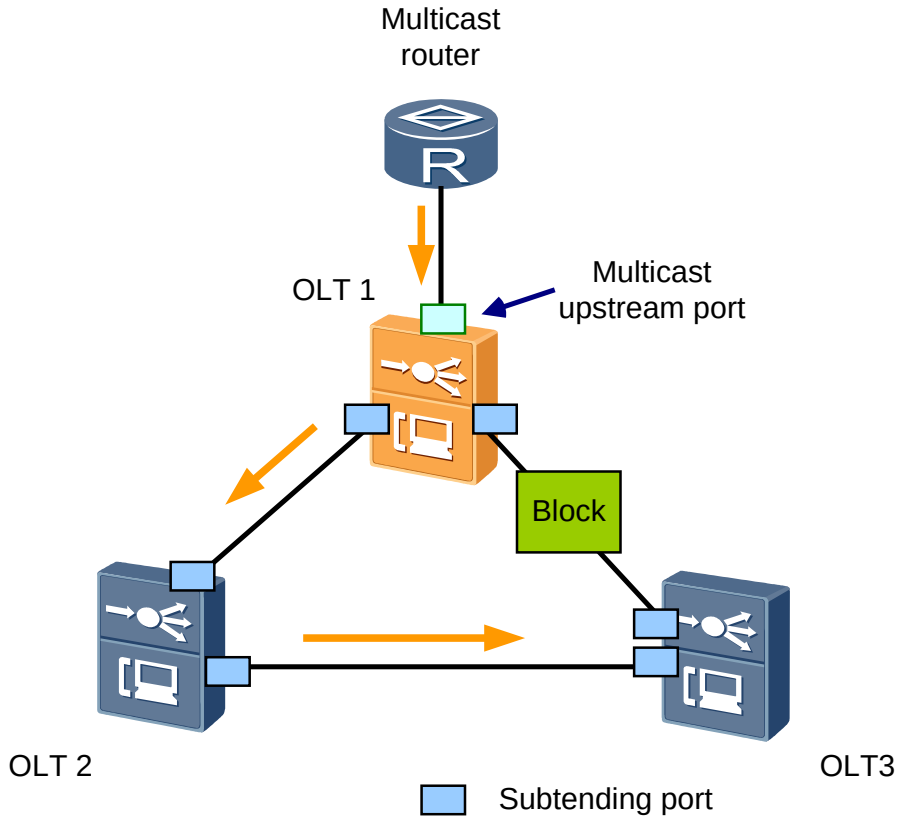


- RRPP (Only the OLT supports this feature)

  The mode of the multicast uplink port needs to be set to RRPP (system-level configuration). In this case, the multicast uplink port of the device does not need to be configured either; instead, the ring uplink port determined dynamically by RRPP serves as the multicast uplink port. The RRPP master node, however, does not need to use the RRPP multicast uplink port mode, but needs to be configured with the correct multicast uplink port and multicast cascading port.

  In addition, the device ports on the ring need to be configured as the multicast cascading ports. The actual multicast downstream ports are determined by IGMP according to the multicast group membership table. The following figure shows the configuration of each role. For details about RRPP, see 19.7 RRPP.

Figure 18-37 Multicast configuration of the RRPP ring network



NOTE

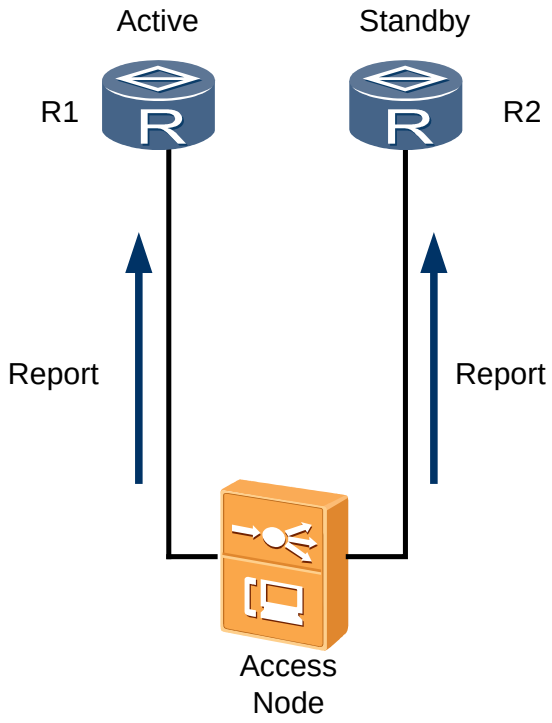
- The RRPP master node must be the injection point of the multicast service.
- When using the ETHA/SPUA/ETHB board for upstream transmission, the device does not support RRPP multicast.
- The device supports only RRPP single ring, and the single ring must be a primary ring.

If a link or device fails, after RRPP selects a backup link, the MVLAN-based IGMP protocol stack immediately sends the new ring uplink port (serving as the multicast uplink port) the join message targeting at the multicast group that the device is interested in. In this way, the multicast service is recovered fast.

Dual-homing of Upstream Ports

Multicast routers 1 and 2 function as the active router and standby router respectively, as shown in the following figure. To ensure fast recovery of the multicast service after a switching, use the IGMP message broadcast function provided by the access node.

Figure 18-38 Upstream port broadcasting IGMP messages



First, set the two access node ports connected to routers 1 and 2 as the multicast upstream ports (the two ports must not be in the same aggregation group or protect group). After this setting, when the access node transmits IGMP messages to router 1, it transmits the same IGMP messages to router 2 at the same time. In this way, router 2 can maintain in real time the same multicast forwarding entry as that of router 1. Once a switching occurs, router 2 can directly obtain the multicast forwarding entry and can ensure fast recovery of multicast service in a shorter time.

Note: If the router supports transfer of the multicast forwarding entry using a proprietary protocol, this can substitute for the upstream port dual-homing function. In this case, add the two access node ports to one aggregation group. Such a function is more commonly used in actual applications.

18.4.4 Advanced Multicast Technologies

Multicast Program Management

Multicast programs can be static programs or dynamic program.

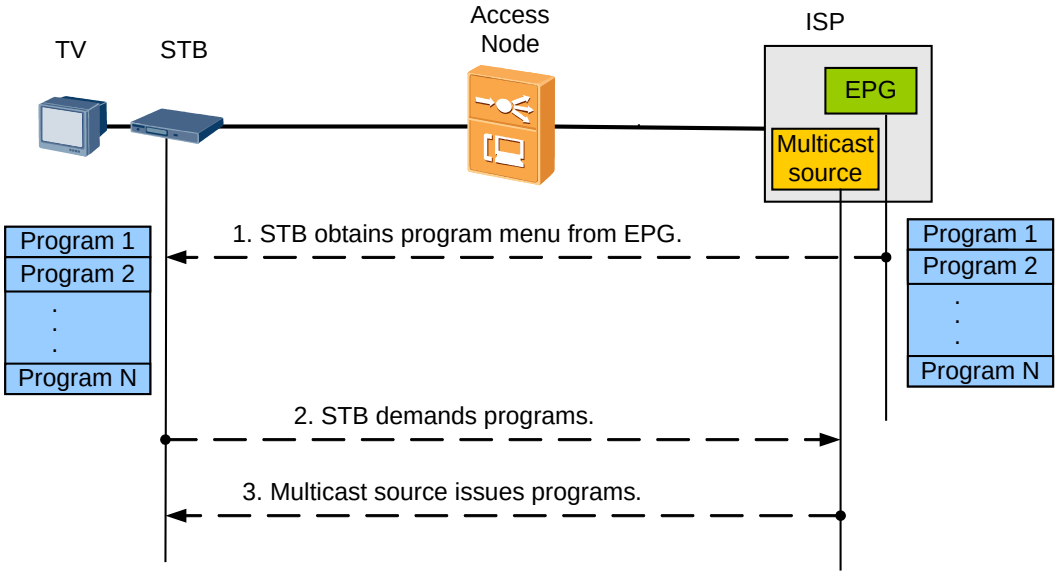
Static Programs

A program list is configured before users can watch these video programs. In this way, controllable multicast is implemented using a rights profile. However, the program list and rights profile must be modified when the video service is modified. The device supports the program host, program prejoin, and multicast bandwidth management functions.

Dynamic Programs

In actual applications, if fine-grained management is not required on the device, dynamic programs can be applied. This avoids maintenance troubles brought by frequent program changes. In this case, program maintenance can be performed uniformly through the Electronic Program Guide (EPG) system.

Figure 18-39 Process of generating dynamic programs



1. After being started, the STB automatically obtains the program menu from the EPG server and provides the menu for the multicast user.
2. When the user orders a program, a corresponding IGMP message is generated and sent to the device. Hence, the program information on the device at this stage is not input by the administrator. Instead, it is dynamically generated in the MVLAN (to which the multicast user belongs) after the multicast group IP address and source IP address are extracted by the device from the real-time IGMP message of the multicast user.
3. The multicast program of the multicast source reaches the STB.

   To prevent the user from using an inappropriate group IP address, a legal multicast address segment can be configured based on MVLAN on the device for dynamic programs. According to the configuration, a multicast program is generated only when the group IP address is within the legal address segment; otherwise, the IGMP message of the user is dropped. Apart from the restriction by the address segment, the number of programs that can be dynamically generated is also controlled by hardware specifications.

   The fine-grained management that is not supported by dynamic programs on the device includes CAC, rights management, multicast preview, and pre-join.

Hot programs are added as static programs and other programs are ordered by users dynamically. This configuration speeds up hot program ordering and shortens channel switching time.

Multicast Rights Management

Rights Management

With the method of configuring different multicast programs on different profiles, package-based rights management can be implemented on the device.

- Rights profile

  The rights to any multicast program can be specified in each rights profile, and each rights profile can be configured with a meaningful name. There are four types of rights:

  Forbidden: It indicates that a multicast user is not allowed to watch or preview a multicast program.

  Preview: It indicates that a multicast user can order a multicast program but is restricted in the watching duration and watching times.

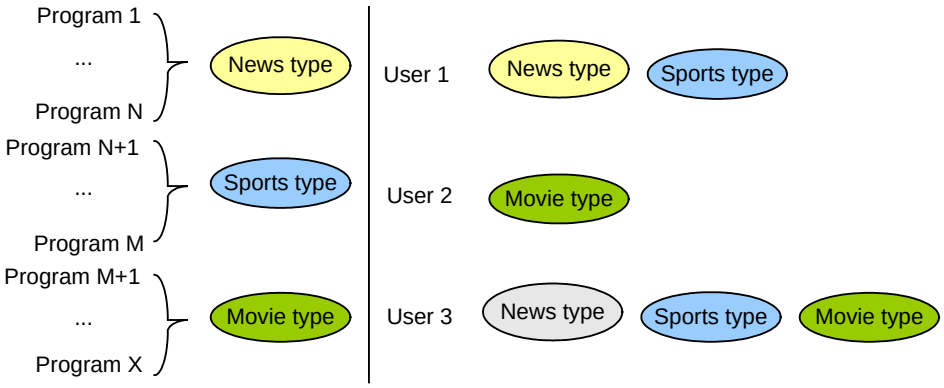
  Watch: It indicates that a multicast user can order a multicast program normally without any restriction.

  Idle: It indicates that a specific right is not assigned to a multicast program yet. It is the default value of the rights profile. The effect of "idle" equals that of "forbidden."

  Carriers can plan the rights profiles according to user-defined rules. Usually, there are three modes of planning.

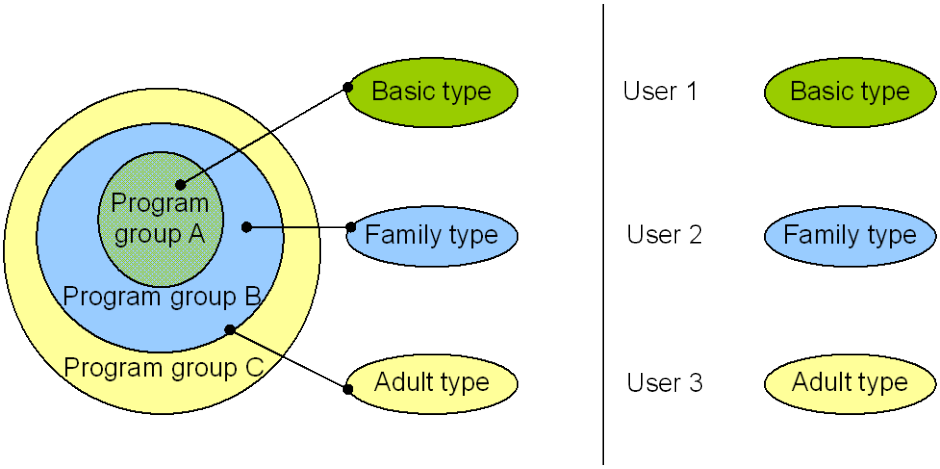
  – The first one is planning by contents, such as planned as the news type, sports type, and movie type. In this case, one multicast program belongs to only one rights profile and the programs of different profiles do not overlap. Therefore, one user is usually bound to multiple profiles. See the following figure.

    Figure 18-40 Planning rights profiles-mode 1

  – The second one is planning different levels by content volumes, such as planned as the basic type, family type, and adult type. In this case, one multicast program may belong to multiple rights profiles and the programs of different profiles may overlap. Therefore, one user is usually bound to only one profile. See the following figure.

Figure 18-41 Planning rights profiles-mode 2



- The third one is a hybrid of the first and second ones and is the most complicated as well as the most flexible mode. In this mode, the programs of different rights profiles may overlap and one user may be bound to multiple profiles. The same program may be configured with different rights in different rights profiles. To ensure that these rights profiles work as expected by the carrier when it comes to a specific program for a specific user, the rights of a program in the rights profiles must be prioritized. It is recommended to plan the priorities before deployment to prevent any incorrect results. The following are examples.

Table 18-8 Priority of rights: forbidden > preview > watch > idle

Rights profile 1	Program 1: watch	User 1	Program 1: forbidden
Rights profile 2	Program 1: forbidden		

Table 18-9 Priority of rights: watch > preview > forbidden > idle

Rights profile 1	Program 1: watch	User 1	Program 1: watch
Rights profile 2	Program 1: forbidden		

- Rights control

  The rights of each multicast user can be configured by the following two steps:

  a. Plan the rights profiles of all multicast programs.

  b. Bind a multicast user to the rights profiles required according to the contents subscribed to by the user.

     The device provides open MIB interfaces to support such operations.

     In addition, there is another method of implementing rights control: by configuring encryption on the head system and the STB. In this way, the carrier does not need to

perform rights management on the device and only needs to enable or disable rights control at the system level or the multicast user level.

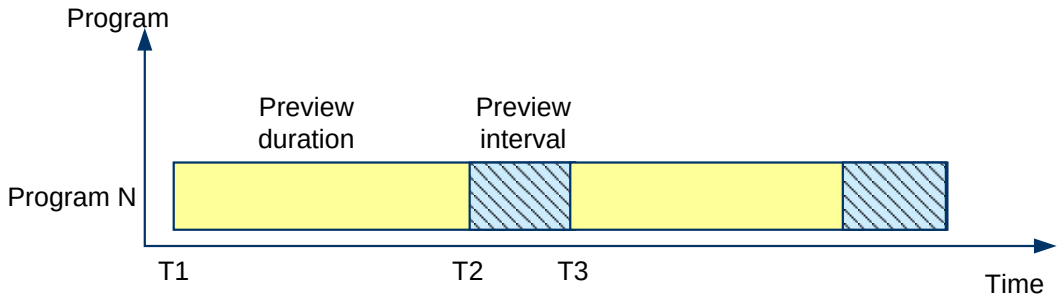
Multicast Preview

By providing the preview of certain special channels to multicast users, carriers may attract more users to subscribe to more programs with the watch right. Preview is usually used as an effective selling method.

The device manages the preview parameters of each multicast program by preview profiles. To be specific, each multicast program can be bound to a preview profile which is configured with preview parameters. Similar programs can be bound to the same preview profile to simplify management.

A preview profile contains three preview parameters.

Figure 18-42 Preview parameters



T1: Start time of first preview

T2: End time of first preview

T3: Start time of second preview

- Preview interval: It is the minimum interval between two previews. The interval is from the end time of the previous preview to the start time of the current preview (from T2 to T3 as shown in the preceding figure). If the interval between the two previews of a user does not reach the specified preview interval, the user is currently not allowed to preview the program. Such a mechanism guards against any "rogue" behavior of users. Without the restriction of the preview interval, a user may keep previewing the same program and is actually "watching" a program without having to pay for the "watch" right.
- Preview times: This parameter specifies how many times a multicast user is allowed to preview the same program during a day. Each time the user leaves a previewed program, the counter increases by 1. When the counter exceeds the maximum value, the further orders of the user for the program will be rejected. In this case, the user's right to the program can be regarded as demoted to "forbidden." However, the preview right can recover the next day.
- Preview duration: This parameter specifies for how long a multicast user is allowed to watch the same program each time. The duration starts from the beginning of the order (from T1 to T2 as shown in the preceding figure). After the duration expires, the user will not be able to receive any data for the multicast program.

For details on how to control the preview of multicast users, see "Rights Management."

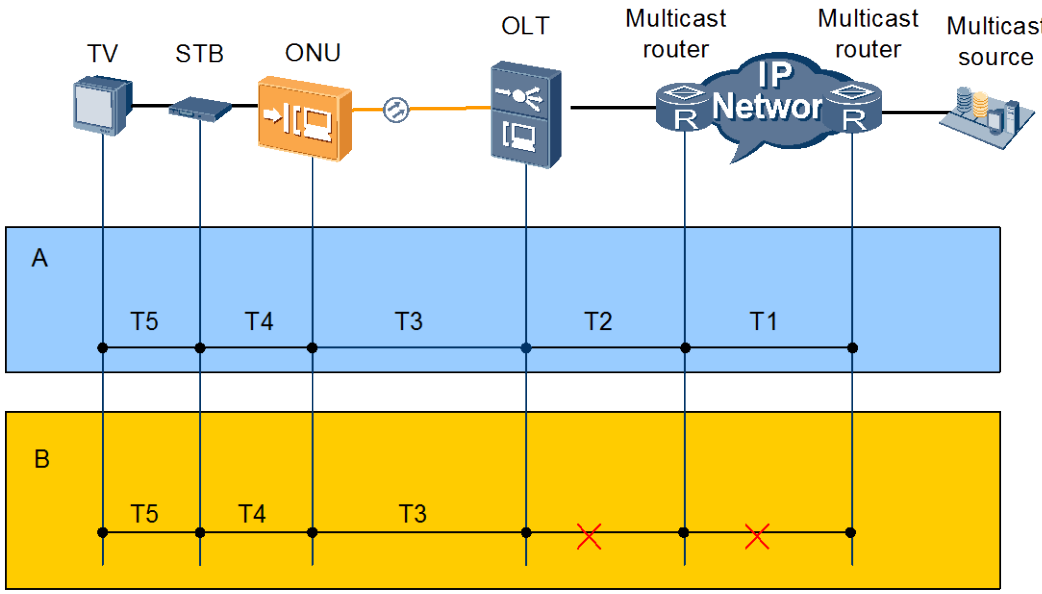
NOTE

For the OLT, boards with centralized multicast services do not support active/standby synchronization of preview data, but the boards with distributed multicast services support so.

Prejoining a Program

The prejoin function is used to shorten the course of channel switching (reduce switching latency), so as to improve users' experience in channel switching. Switching latency comprises the processing consumption in each segment of a network, as shown in the following figure. With the prejoin function enabled, the network-side processing consumption (T1+T2) equals 0.

Figure 18-43 E2E multicast switching latency



A: Processing latency of non-prejoin program = T1 + T2 + T3 + T4 + T5

B: Processing latency of prejoin program = T3 + T4 + T5

The prejoin function applies to the IGMP proxy scenario. It equals that there is always online users for a program.

- The flow of prejoining a program is the same as the flow of joining a program normally. For details, see section 1.6.4. Once a multicast stream is successfully demanded, it is transmitted to the DSLAM.
- In the flow of leaving a prejoined program, compared with the flow of leaving a normally joined program, the DSLAM does not transmit the leave message to the multicast router even when the last multicast user leaves the program.
- In the flow of querying a prejoined program, compared with the flow of querying a normally joined program, the DSLAM responds to the multicast router's query as required by the protocol regardless of whether or not the multicast group membership table of the program contains a multicast user.

From above all, viewed from the router, there are always online users for a prejoined program.

The prejoin function can be set for a program. In general, set the prejoin function for the program that is most commonly demanded by users. A dynamic program does not support the prejoin function.

Multicast CAC

CAC is the short form for call admission control. Here, it means controlling the setup of IGMP sessions. If an IGMP session fails to be set up, a multicast user will fail to receive the multicast program ordered.

In a broad sense, implementing CAC requires implementing the first-level control in the system. Currently, system control includes the following:

- Anti-DoS attack. The rate of IGMP messages sent from the user side must not exceed the specified value in the system. Otherwise, the system will regard that a DoS attack occurs and drops the messages. Such a protection method applies not only to IGMP messages, but also to control packets such as DHCP and PPPoE packets. For details, see "Anti-DoS Attack."
- Anti-IP spoofing. When this function is enabled, the user must obtain a legal IP address through DHCP before ordering any program. Only the IGMP messages using the legal IP address as their source IP address will be accepted by the system; otherwise, the messages will be regarded as coming from unauthorized users and will be dropped by the system. For details, see "Anti-IP Spoofing."

NOTE
Only the centrally-controlled multicast supports this feature.

- Broadband message overload. When a service traffic burst occurs, the system resources may not be able to support all services. Then, the system will drop certain messages according to specified policies to ensure that the services with a higher priority are not affected. In this case, IGMP messages may be "sacrificed" to reduce the system load. For details, see "Broadband Message Overload."

After the first-level control in the system comes the multicast first-level control, which includes the following:

- Concurrent number of programs of a multicast user. This parameter specifies how many channels a multicast user is allowed to order at the same time. The parameter can be configured based on multicast user.
- Rights control. For details, see "Multicast Rights Management."
- Bandwidth check. Though the system supports QoS control on various types of traffic, packet loss (drop by priority or tail drop) may still occur when the transmission bandwidth is overloaded. However, due to the real-time and non-retransmittable properties of multicast programs, postmortem QoS will directly cause pixelation to the programs with packet loss (not only to newly ordered programs). Hence, the requirements of IPTV for high-quality experience are not met. Bandwidth check enables the system to control a newly ordered program beforehand. In this way, the system can ensure that the programs that have been ordered enjoy sufficient bandwidth and will not be affected by the new program. With bandwidth check, only the newly ordered program is affected (if bandwidth is insufficient, the user will not be able to watch the newly ordered program).

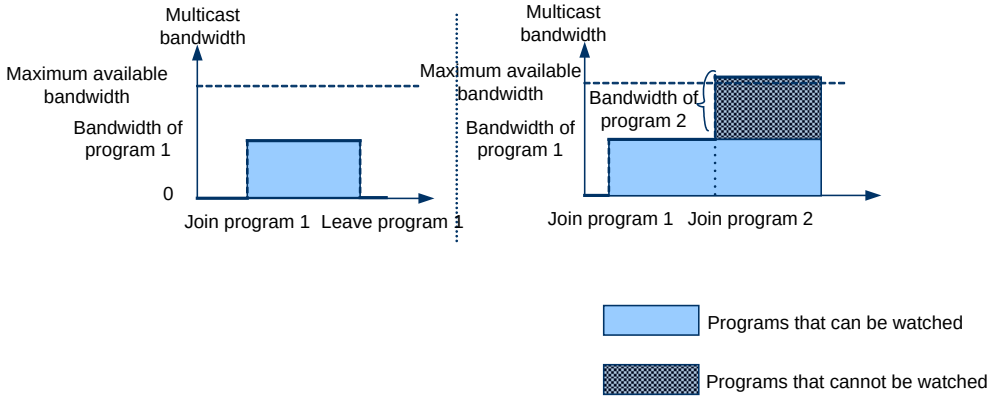
CAC can be classified into three types according to different control points and methods.

- Multicast user bandwidth CAC

  First, each pre-configured program is configured with bandwidth. The bandwidth is configured with reference to the video bit streams, and the margin of packet encapsulation and network transmission jitter; if possible, actually tested network traffic

can also be used, as a better reference. Then, each multicast user is configured with available bandwidth. The available bandwidth is configured with reference to the actual line bandwidth or the planning of service provisioning. Hence, when receiving the first IGMP join message of a program, the device subtracts the bandwidth occupied by the program from the available bandwidth of the user. If the remainder is smaller than 0, the device rejects the order request of the user. When receiving an IGMP leave message of a program, the device returns the bandwidth occupied by the program to the available bandwidth of the user. The time of returning is the time when the device stops forwarding multicast data. That is, the program is not ordered by any end user of the terminal.

Figure 18-44 Multicast user bandwidth CAC



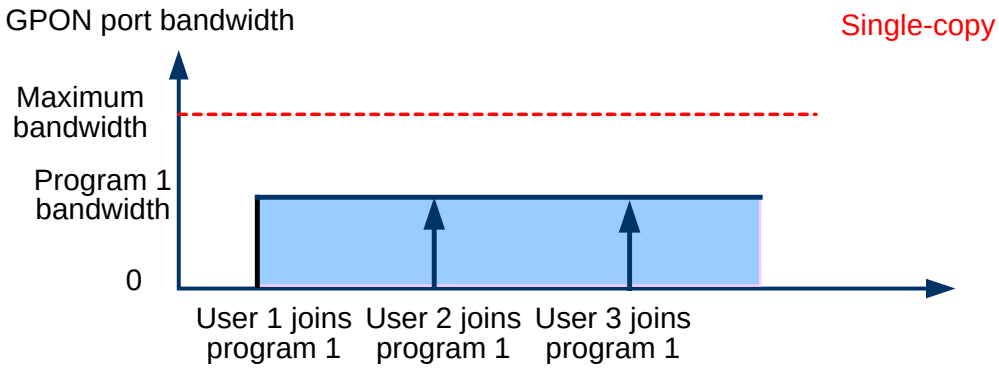
Multicast user bandwidth CAC can be configured at the system level or at the multicast user level.

- GPON port bandwidth CAC [OLT]

  GPON single-copy duplication function (default configuration): Under the same GPON port, even if multiple multicast users order the same multicast program, the multicast data is duplicated only once and sent to corresponding multicast users through the downstream multicast channel. Therefore, this function ensures that the downstream multicast bandwidth does not overflow the downstream line bandwidth of the GPON port.

  To do so, the operator first needs to configure bandwidth for each pre-configured program (see "Multicast user bandwidth CAC"), and then allocates the available bandwidth for each GPON port (depending on the actual line bandwidth or the service provisioning plan). In this way, after receiving the first IGMP join message of the program, the device deducts the bandwidth of the corresponding program from the remaining bandwidth of the GPON port. If the deduction result is smaller than 0, the device rejects the order of the user. After receiving an IGMP leave message, the device returns the bandwidth of the corresponding program to the GPON port (the moment of returning is when the forwarding of multicast data is stopped, that is, no multicast user under the GPON port orders this program).

Figure 18-45 GPON port bandwidth CAC



This function can be configured at the system level or GPON port level. It can be used together with the multicast user bandwidth CAC.

- GPON port ANCP bandwidth CAC [OLT]

  Generally, the IPTV service includes unicast stream and multicast stream, corresponding to VOD service and TV service respectively. By using ANCP, this bandwidth CAC function can be used on the RACS and VOD servers to implement bandwidth CAC (not only multicast bandwidth CAC) on all IPTV traffic streams. (For the ANCP principle, see "ANCP Feature.")

  This bandwidth CAC function can be configured at the system level. It cannot be used with the previous two bandwidth CAC functions.

Multicast QoS

This topic describes the quality of service (QoS) features dedicated to the multicast service of the OLT. For more information about QoS (including traffic classification, traffic policing, ACL policy, and congestion avoidance and management), see the 14 QoS.

Priority Processing of IGMP Packets

The device supports processing of only the 802.1p priority (CoS) of IGMP packets.

Table 18-10 Priority processing in IGMP proxy/snooping

Cscading Mode	Upstream	Downstream
Multicast user/xPON cascading	Based on MVLAN.	Traffic classification methods: • VLAN: by the 802.1p priority specified in the traffic profile. • VLAN+encapsulation type: by the 802.1p priority specified in the traffic profile. • VLAN+802.1p priority: by the 802.1p priority specified by traffic classification.

Priority Processing of Multicast Traffic Streams

The device supports processing of only the 802.1p priority (CoS) of the multicast traffic streams.

Table 18-11 Downstream multicast priority processing

Cscading Mode	Pre-configured Program	Dynamic Program
Multicast user/xPON cascading	Traffic classification methods: • VLAN: by the 802.1p priority specified in the traffic profile. • VLAN+encapsulation type: by the 802.1p priority specified in the traffic profile. • VLAN+802.1p priority: by the 802.1p priority specified by traffic classification.	

GPON ONT Multicast

The GPON end-to-end multicast service requires the cooperation of the ONT. The following points must be noted:

- VLAN translation

  If the carrier plans the home gateway at the user's house, generally, the VLAN of the IPTV service (also called C-VLAN) needs to be planned. Because the OLT does not directly support translation of the MVLAN, the operator can configure VLAN translation on the ONT to meet the planning requirement (the OLT provides the corresponding CLI and the configuration can be issued to the ONT through OMCI). The MVLAN can be translated in three ways: transparently transmitted, translated to untagged, and translated to a specified VLAN.

- Controllable multicast

  In single-copy duplication, GPON downstream multicast programs are broadcast. Assume the following condition: After an authorized multicast user orders a program, all users under the GPON port can receive this program. Therefore, to implement complete rights control on the access device, the OLT must configure the ONT to work in the "dynamic controllable" mode. In this way, the multicast filtering table (white list) on the ONT is issued by the OLT after multicast control checking. If a downstream multicast program is not in the multicast filtering table, the ONT cannot receive this multicast program.

  If the ONT is configured to work in the "IGMP snooping" mode, the multicast filtering table on the ONT is completely maintained by the ONT. In this case, multicast program rights management is generally implemented by the encryption system of the IPTV platform.

Multi-instance Multicast

With the increased use of open networks, carriers' networks need to provide independent multicast domains for different multicast ISPs so that different ISPs do not interfere with each other. Independent multicast domains can be implemented on the management plane, control plane, and forwarding plane by planning different MVLANs on the device.

For OLT

- Management plane (ASMSSM or ASM ONLY)

  Within each MVLAN, the multicast programs to be provisioned, and the multicast upstream ports and multicast users involved can be configured for each ISP. Here, the multicast programs need to be noted. To ensure that each ISP can plan multicast programs independently, the multicast program triplet (MVLAN, source IP address, and multicast IP address) needs to observe the following rules:

  – If two GIPs are mapped to the same GMAC (for details on the mapping method, see "Basic Concepts"), the two GIPs are regarded as the same GIP.

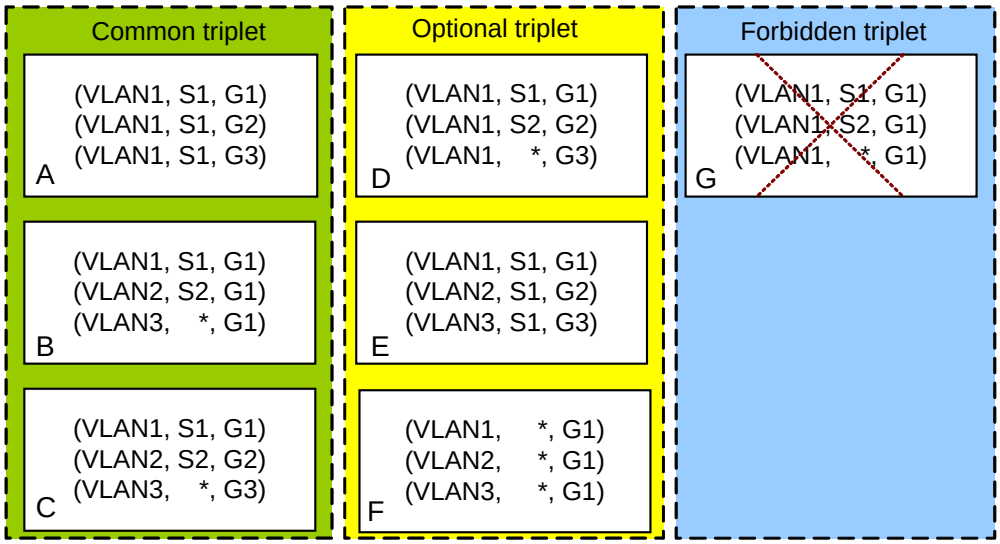
  NOTE
  If the device does not have boards that support only VLAN+GMAC forwarding chip, this rule may not be considered; otherwise, carriers need to follow this rule in planning.

  – To ensure the uniqueness of a multicast forwarding entry on the forwarding plane, (MVLAN, multicast IP address) must be unique.

  – Especially, in the case of an IGMPv2 message or an IGMPv3 message in the ASM mode, the multicast source IP address equals any value (usually represented as * or as any). In this case, only the second rule needs to be observed.

  Use section G in the following figure as an example. According to the second rule, (MVLAN, multicast IP address) must be unique, but (VLAN1, G1) in section G is not unique. Therefore, configuring or generating the entries in section G is not allowed. The entries in the other sections in the following figure can also be judged by the rules described above.

  Figure 18-46 Multicast program triplet (ASMSSM or ASM ONLY)



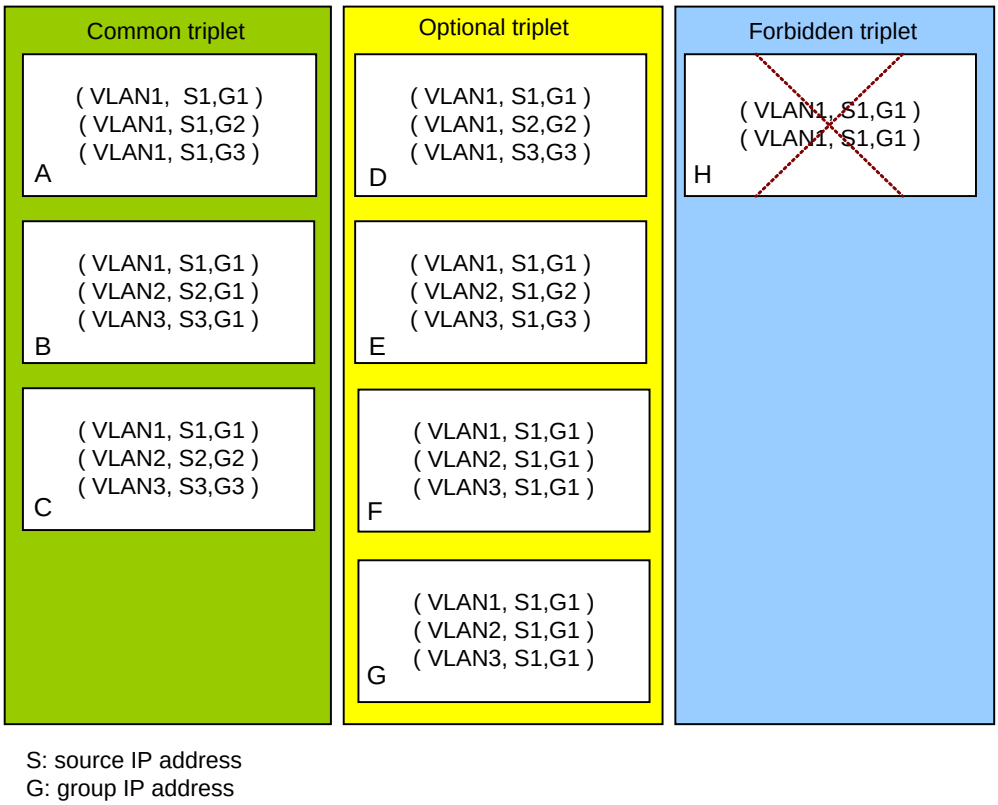
- Management plane (SSM ONLY)

  Within each MVLAN, the multicast programs to be provisioned, and the multicast upstream ports and multicast users involved can be configured for each ISP. Here, the multicast programs need to be noted. To ensure that each ISP can plan multicast programs independently, the multicast program triplet (MVLAN, source IP address, and multicast IP address) needs to meet the following rules:

- If two source IP addresses have the same least significant 20 bits, these two source IP addresses are considered the same source IP address. Carriers need to have proper planning to avoid this.
- To ensure the uniqueness of a multicast forwarding entry on the forwarding plane, (MVLAN, multicast IP address, and source IP address) must be unique.

Use section H in the following figure as an example. According to the second rule, (MVLAN, multicast IP address, and source IP address) must be unique, but (VLAN1, G1, and S1) in section G is not unique. Therefore, the entries in the other sections in the following figure can also be judged by the rules described above.

Figure 18-47 Multicast program triplet (SSM ONLY)



- Control plane

  On the network side, each MVLAN has an independent IGMP protocol stack. Each ISP can select the protocol version, message priority, and IGMP proxy or IGMP snooping. On the user side, each multicast user has an independent IGMP protocol stack and is not affected by other multicast users. Table 18-12 describes the IGMP message processing policies in different group filter modes.

Table 18-12 IGMP message processing policies

Group Filter Mode	IGMPv2	IGMPv3 Without Source	IGMPv3 with Source
ASMSSM	Processed	Processed	Processed

Group Filter Mode	IGMPv2	IGMPv3 Without Source	IGMPv3 with Source
ASM ONLY	Processed	Processed	Dropped
SSM ONLY	Dropped	Dropped	Processed

- Forwarding plane

  On the forwarding plane, multicast forwarding tables with the group filter mode use ASMSSM or ASM ONLY use the MVLAN and multicast IP address together as indexes (VLAN+GMAC as the index for boards that do not support VLAN+GIP). This ensures that different MVLANs do not interfere with each other. Multicast forwarding tables with the group filter mode SSM ONLY use MVLAN, multicast IP address, and source IP address together as indexes. For details, see "Forwarding Framework on the Device." For the control board, implementing QoS scheduling on the traffic of different MVLANs on the same port on the service board equals unicasting. For details, see "QoS."

NOTICE

1. In most situations, the MVLAN is different from the S-VLAN of the traffic stream to which the multicast user belongs. This S-VLAN can be used in the IPTV VoD service.
2. VLANs in S-VLAN+C-VLAN forwarding mode are used for transmission over point-to-point channels and are not applicable to multicast duplication and forwarding. Therefore, the S-VLAN in S-VLAN+C-VLAN forwarding mode cannot be used as an MVLAN.
3. When the group filter mode is ASMSSM, binding a multicast user to multiple MVLANs is not recommended. If a home has subscribed to multiple ISP programs, configure different multicast users for different MVLANs to ensure that one-to-one mapping between MVLANs and multicast users.
4. When the group filter mode is ASM ONLY or SSM ONLY, binding a multicast user to multiple MVLANs is not supported.

Charging Mode

For multicast services, carries or ISPs usually adopt two charging modes:

- Fixed charging. In this mode, programs are divided into different packages. The user needs to pay a fixed amount of fee for each package in a fixed period (such as by the year or by the month). This charging mode does not restrict the multicast user in the order count or the ordered volume of traffic.
- Pay per view (PPV): In this mode, the user is charged according to the order count of different programs.

In the first charging mode, since it is fixed charging, the charging does not relate to the behavior of the multicast user. Therefore, the first mode is supported by the device inherently and does not require additional functions from the device.

In the second mode, the device records the order behavior of each multicast user or each multicast terminal identified by the MAC address, and provides the behavior information in

the form of a call detail record (CDR) to the accounting system for settling an account. The complete configuration of the CDR function consists of three steps:

1. Enabling the logging function. The function can be configured at the multicast user level, multicast program level (configurable for pre-configured programs, and defaulted to **enable** for dynamic programs), or system level. When a user finishes a complete watch behavior (from the program order starting to ending), or when the user fails to order a program because of failing to pass the multicast CAC, a log is generated.

NOTE

When the logs reach the maximum capacity, new logs will overwrite old ones. Therefore, to prevent heavy consumption of log resources in the case where the user quickly browses through channels, the device supports a configurable flag time for generating logs. If the multicast user watches a channel for a duration shorter than the flag time, the device does not generate a log. On the contrary, to timely log the users who stay online for a long time, the device supports the configuration of another time value (if the value is 0, the log will not be generated). When a user stays online for longer than the preset value, the device automatically generates a log.

2. Configuring the file server. The operator needs to select a CDR transfer file. Available options are TFTP/FTP/SFTP. Also, the operator needs to set the IP addresses of the primary and secondary servers.

NOTE

SFTP is recommended.

3. Enabling the CDR functions (at system level). After the CDR function is enabled, the device automatically integrates the logs that need to be reported into a text file and transfers the file to the server when either of the following conditions is met: when the reporting interval expires, or when the number of logs reaches the reporting threshold.

   There are two types of CDR but they have the same format of the text file, as showed in Figure 18-48. The format of the text file name is different from the detailed format of a CRD item.

Figure 18-48 Format of the text file

   - Recording the order behavior of each multicast user

     The format of the text file name is **HWCDR**-*host name*-*YYYYMMDDHHMMSS*.**txt**.

Table 18-13 Detailed format of a CDR item

ID	Field Name	Specification	Commentary
0	TAG	3 Bytes	Fixed as "Log". "Log" is the module name which generate syslog
1	SN	0..5 Bytes	Using a 16-bit variable to record. The maximum value is "65535" which

ID	Field Name	Specification	Commentary
			occupies 5 bytes.
2	FrameSlotPortGemport	5..13 Bytes	F/S/P/GemPort for GPON user
	FrameSlotPortFlow	5..14 Bytes	F/S/P/FlowID for other type user
3	ProgramIP	0..15 Bytes	Sample: 239.1.1.1
4	OperMode	0..1 Bytes	0-Watch; 1-Preview; 2-No Right; Other is invalid
5	StartDate	0..18 Bytes	YYYY-MM-DD HH:MM:SS
6	EndDate	0..18 Bytes	YYYY-MM-DD HH:MM:SS
7	ProgramName	0..16 Bytes	Sample: cctv1. If the program does not exist, this parameter is "No-Name".
8	ProgramSrcIP	0..15 Bytes	Sample: 192.168.1.1. If the IP address is invalid, this parameter is "*".
9	Reason	1..2 Bytes	Syslog generation reason: 11: User's online time is too long 0: User leave.

- Recording the order behavior of each multicast user or each multicast terminal identified by the MAC address

  The format of the text file name is **BTVCDR**-*host name-X-YYYYMMDDHHMMSS****.txt**.

NOTE

When the value of X is A, automatic reporting is enabled. When the value of X is M, manual reporting is required. *** indicates millisecond.

Table 18-14 Detailed format of a CDR item

ID	Field Name	Specification	Commentary
1	Record type	1..2 Bytes	• 0: Successful channel zap • 1: Channel timeout by general query or group-specific query • 2: Successful preview channel zap • 3: Preview channel timeout by general query or group-specific query • 4: It is automatically generated every N hours even when there is no channel (including preview channel) zap. • 5: Join a channel with no access right • 6: Join a channel with preview access

ID	Field Name	Specification	Commentary
			right (but with preview limit exceeded)
2	System name	1..50 Bytes	The name of the OLT/DSLAM.
3	IGMP user	7..14 Bytes	Frame ID/Slot ID/Port ID/Service port Index Sample: 0/2/0/233
4	Terminal IP	7..15 Bytes	Zero-length string means IP is unknown. Sample: 192.168.0.1
5	Terminal MAC	17 Bytes	Zero-length string means MAC is unknown. Sample: 00-1B-21-B4-0B-EE
6	Multicast VLAN	1..4 Bytes	Multicast VLAN of the program Sample: 200
7	Program Group IP	9..15 Bytes	Zero-length string means source IP of the program is not configured. Sample: 10.1.1.1
8	Program Source IP	0..15 Bytes	Sample: 192.168.1.1. If the IP address is invalid, this parameter is "*".
9	Start viewing time	19 Bytes	YYYY-MM-DD HH:MM:SS{+|-}hh:mm [DST]. Sample: 2011-08-27 09:30:20+08:00
10	Time of record generation	19 Bytes	YYYY-MM-DD HH:MM:SS{+|-}hh:mm [DST] Sample: 2011-08-27 10:30:20+08:00
11	Program Name	1..16 Bytes	Sample: cctv1
12	Duration viewed	1..n Bytes	Viewing duration in seconds. Sample: 30

Double-VLAN Tag Multicast

Double-tag multicast specially refers to the number of VLAN tags of packets carried by the network-side multicast of the access device. In actual multicast networks, generally only one VLAN tag is used for multicast. The reasons for using double-tag multicast are as follows:

- The unicast application is limited by VLAN number supported by device and more and more applications adopt the double-tag planning. In addition, because the convergence switches of some vendors interconnected with the access device do not support transmission of single-tag packets and double-tag packets on the same physical link, the multicast has to adopt the double-tag mode.

- To use the VLAN planning like that of unicast in a unified manner. For example, outer VLAN tags indicate different ISPs and inner VLAN tags indicate different services.

The device can be configured whether to use the double-tag multicast function.

- IGMP control message

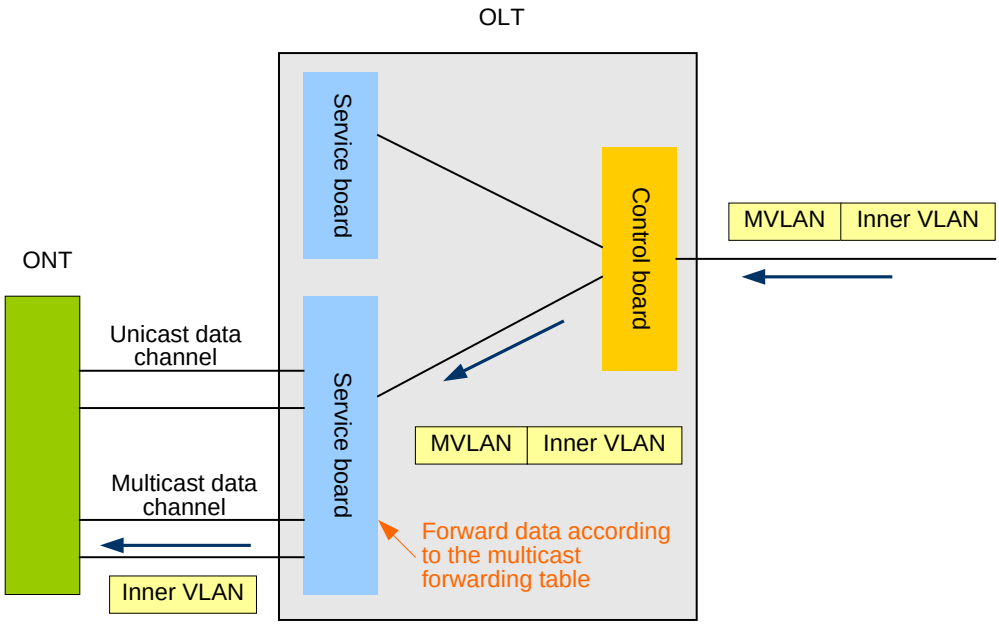
  Double-tag multicast and single-tag multicast have the same processing flow for IGMP control messages (see "Join Flow"). Difference between them: On the network side, the transmitted and received IGMP messages in single-tag multicast have one tag whereas the transmitted and received IGMP messages in double-tag multicast have two tags The outer VLAN tag in double-tag multicast is the MVLAN to which the program belongs and the inner VLAN tag can be configured based on MVLAN (adopting the "easy in strict out" principle, the device does not check the inner tag and configuration consistency of received IGMP messages).The inner tag priority and outer tag priority of IGMP packets can be configured separately based on MVLAN, but the inner tag priority must be the same as the outer tag priority in the same MVLAN and cannot be configured separately. If interconnection with a device of non-0x8100 TPID is required, the MVLAN is configured to a stacking VLAN and then TPIDs of inner tag and outer tag are configured globally. (Currently, port-based TPID configuration on the SPUA upstream board is not supported.)
- Multicast data hardware forwarding

  The data forwarding flow of double-tag multicast streams varies according to the GPON duplication mode. If the multicast duplication mode is multicast (single-copy duplication), multicast streams are forwarded on the control board and service board according to the multicast forwarding entry. The VLAN carried by the multicast streams forwarded to the ONT is the CVLAN. Because of the multicast duplication feature, in this scenario, the inner VLAN is often directly defined as the MVLAN (In addition, the device currently supports only this scenario).

NOTE

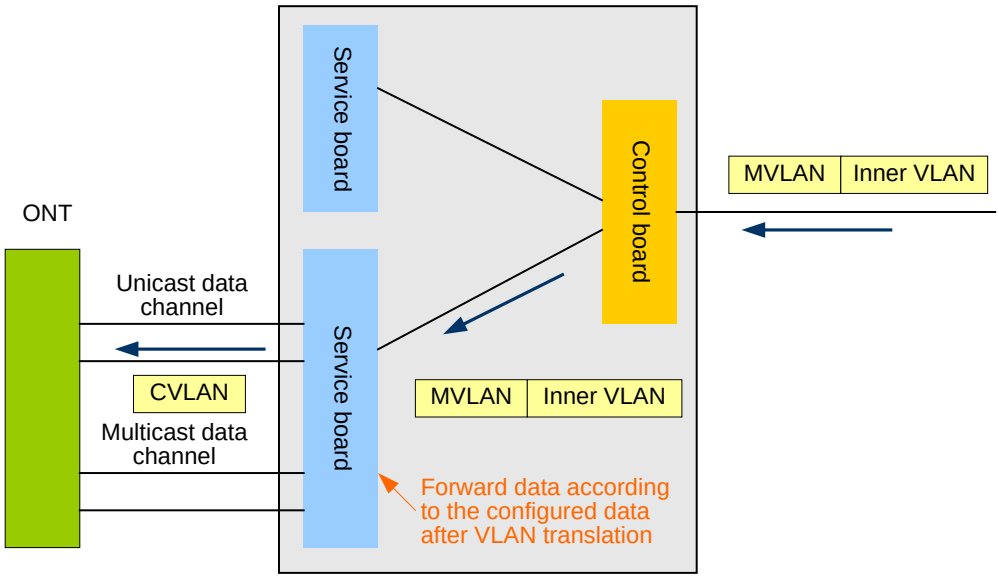
This scenario is applicable only when the TPID is 0x8100.

Figure 18-49 Double-tag multicast hardware forwarding (single-copy duplication)



If the multicast duplication mode is unicast (multi-copy duplication), multicast streams are forwarded on the control board according to the multicast forwarding entry and on the service board according to the configured service port. The VLAN carried by the multicast streams forwarded to the ONT is the CVLAN. In this scenario, the inner VLAN is often defined as the CVLAN.

Figure 18-50 Double-tag multicast hardware forwarding (multi-copy duplication)



Transparent Transmission of Multicast Data

- Transparent transmission of multicast data for private-line users
  - Transparent transmission of IGMP messages

    In the upstream direction, the service board determines whether to transparently transmit the received IGMP messages according to the VLAN-level IGMP transparent transmission policy and service-port-level transparent transmission policy. If transparent transmission is allowed, IGMP messages are added with the SVLAN tag on the service board and then are transmitted to the control board. After arriving at the control board, IGMP messages are broadcast by the control board within the SVLAN. If transparent transmission is not allowed, IGMP messages are dropped.

    In the downstream direction, the IGMP messages transmitted from the network side are broadcast by the control board within the SVLAN and after they arrive at the service board, the service board determines whether to transparently transmit them according to the VLAN-level IGMP transparent transmission policy and service-port-level transparent transmission policy. If transparent transmission is allowed, the service board translates the SVLAN tag to the CLAN tag according to the configuration of the traffic stream and then transmits the messages to users. If transparent transmission is not allowed, IGMP messages are dropped.
  - Transparent transmission of unknown multicast data

    In the upstream direction, the service board determines whether to transparently transmit the received unknown multicast data according to the VLAN-level IGMP transparent transmission policy and service-port-level transparent transmission policy. If transparent transmission is allowed, IGMP messages are added with the

SVLAN tag on the service board and then are transmitted to the control board. After arriving at the control board, IGMP messages are broadcast by the control board within the SVLAN. If transparent transmission is not allowed, IGMP messages are dropped.

In the downstream direction, the unknown multicast data transmitted from the network side is broadcast by the control board within the SVLAN and after the data arrives at the service board, the service board determines whether to transparently transmit the data according to the VLAN-level IGMP transparent transmission policy and service-port-level transparent transmission policy. If transparent transmission is allowed, the service board translates the SVLAN tag to the CLAN tag according to the configuration of the traffic stream and then transmits the data to users. If transparent transmission is not allowed, unknown multicast data is dropped.

NOTE

To prevent the multicast data of the multicast user provisioned with multicast service from being transmitted to the upstream unauthorized multicast sources, make sure that the policy of transmitting unknown multicast data is set to drop. The transparent transmission policies of unknown multicast traffic have the switches of two levels on a service board: the VLAN level and the service port level. When the two switches are both set to transparent transmission, the policy is transparent transmission. When either of the two switches is set to drop, the policy is drop. (Only transparent transmission is supported for connection-oriented traffic and the policy is not configurable in this case.)

- Co-existence of IPTV service and transparent transmission of multicast data
  - Multi-service-port solution

    IPTV service and multicast transparent transmission service are carried on two service ports, and the SVLAN of the service port that carries multicast transparent transmission service must not be the MVLAN.

    The service port that carries IPTV service processes the received IGMP messages following the flow of processing IPTV service, and forwards the multicast data according to the multicast forwarding entry.

    The service port that carries multicast transparent transmission service transparently transmits or drops the received IGMP messages according to the IGMP transparent transmission policy of the traffic stream, and transmits or drops the received unknown multicast data according to the unknown multicast transparent transmission policy of the traffic stream.
  - Single-service-port solution

    IPTV service and multicast transparent transmission service are carried on one service port, whose SVLAN must not be the MVLAN.

    When the Access Node receives upstream IGMP messages, it matches the multicast group address in IGMP messages to the programs in the MVLAN. If the group address successfully matches a program, the Access Node processes the messages as IPTV service. If the group address fails to match any program, the Access Node determines whether to transparently transmit the messages according to the IGMP transparent transmission policy of the SVLAN and service port. The Access Node transparently transmits the messages only when the IGMP transparent transmission policy is enabled for both the SVLAN and service port.

    If the Access Node receives downstream IGMP messages that carry the MVLAN tag, the Access Node processes the IGMP messages as IPTV service. If the messages carry the SVLAN tag, the Access Node forwards them according to the IGMP transparent transmission policy of the SVLAN and service port. The Access Node transparently transmits the messages only when the IGMP transparent transmission policy is enabled for both the SVLAN and service port.

If the Access Node receives the multicast data of IPTV service, the Access Node forwards the multicast data according to the multicast forwarding entry. If the multicast data is unknown, the Access Node forwards the data according to the unknown multicast transparent transmission policy of the VLAN and service port. The Access Node transparently transmits the data only when the unknown multicast transparent transmission policy is enabled for both the SVLAN and service port.

18.5 IPv6 Multicast

This topic describes the aspects unique to IPv6 multicast and the differences between IPv4 multicast and IPv6 multicast.

18.5.1 Introduction to IPv6 Multicast

Overview

As a substitution of IPv4, IPv6 uses a 128-bit address structure to resolve IP address shortage issues and also optimizes some features. The difference between IPv6-based multicast and IPv4-based multicast lies in the significant increase of IP addresses. Other functions of the two types of multicast, such as group member management, multicast packet forwarding, and multicast routing setup, are basically the same.

Purpose

After an IPv4 network evolves to an IPv6 network, IPv6 multicast technologies can be used to provide carriers with a comprehensive set of IPv6 video services, such as live TV and near video on demand (NVoD).

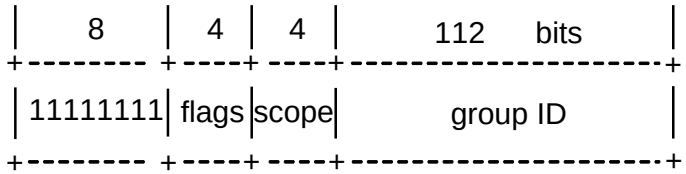
Using IPv6 multicast technologies, the network device can manage, control, and forward IPv6 video services and in this way meets carriers' requirements for provisioning IPv6 video services.

18.5.2 Principle

IPv6 Multicast Address

Figure 18-51 shows the format of an IPv6 multicast address as defined in RFC4291.

Figure 18-51 Format of an IPv6 multicast address



The most significant eight bits of an IPv6 multicast address are consistently 11111111; therefore, an IPv6 multicast address always starts with FF.

Flags is a 4-bit field that indicates four flags set on the multicast address. The four flags are defined as |0|R|P|T|. The highest-order flag is always set to 0. The meanings of the other flags are as follows:

Table 18-15 Meaning of the flags Field

Name of the Bit	Value	Description
R	0	Indicates a multicast address that does not embed the unicast address on the rendezvous point (RP) of the multicast group.
	1	Indicates a multicast address that embeds the unicast address on the RP of the multicast group.
P	0	Indicates a multicast address that is not assigned based on the unicast prefix.
	1	Indicates a unicast-prefix-based multicast address.
T	0	Indicates a permanently assigned multicast address.
	1	Indicates a transient (not permanently assigned) multicast address. **NOTE** Multicast services in this document all use transient multicast addresses. Therefore, the T bit must be set to 1.

Scope is a 4-bit field that indicates the multicast scope. This field is used to signify the scope of a multicast group, that is, whether the multicast group contains nodes of the same local network, site, or organization, or nodes from the global address space. The values of this field are defined as follows:

Table 18-16 Meaning of the Scope Field

Value of the Scope Field	Description
0	Reserved
1	Node/Interface-local scope
2	Link-local scope
3	Reserved
4	Admin-local scope
5	Site-local scope
8	Organization-local scope

Value of the Scope Field	Description
E	Global scope
F	Reserved
6, 7, 9, and A-D	Unassigned

Group ID identifies the multicast group.

Ethernet Multicast MAC Address

RFC2464 defines a set of rules for mapping IPv6 multicast addresses to MAC addresses. An IPv6 address is mapped to the MAC address 3333.XXXX.XXXX, with the 32-bit XXXX.XXXX copied from the least significant 32 bits of the IPv6 address.

Protocol Interoperation

IPv6-based group members are managed using Multicast Listener Discovery (MLD).

MLD is a sub-protocol of Internet Control Message Protocol version 6 (ICMPv6). MLD establishes and maintains the multicast group membership between a user host and its immediately neighboring multicast router. MLD can be regarded as the Internet Group Management Protocol (IGMP) in IPv6, as MLD and IPv6 IGMP have similar implementations.

MLD has two versions: MLDv1 and MLDv2. MLDv2 is fully compatible with MLDv1 and includes all basic concepts of MLDv1.

- MLDv1 (defined in RFC2710)

  MLDv1 is derived from IGMPv2 and supports any-source multicast (ASM), but supports source-specific multicast (SSM) through SSM mapping.
- MLDv2 (defined in RFC3810)

  MLDv2 is a translation of IGMPv3 for IPv6 semantics and supports ASM and SSM. Compared with MLDv1 (defined in RFC2710), MLDv2 has the following improvements:
  - Batch report. The destination IP address in the IPv6 header of a report message is always filled in as FF02:0:0:0:0:0:0:16. In addition, the MLD payload can carry multiple group records, reducing the number of report messages between devices. In an MLDv1 report message, the destination IP address must be filled in as the corresponding group IP address, so one MLDv1 report message cannot carry multiple group records.
  - Longer maximum response time for a query message. The maximum response time for a query message is extended from 65.5s (in MLDv1) to 8387.5s (5000s supported by Huawei devices). Therefore, MLDv2 is applicable to large-scale networks.
  - Source filter. With the source filter function, the host can receive or not receive the multicast data carrying the IP address of a specified multicast source. This function enables the device to better implement source-specific multicast (SSM) and support the multi-ISP scenario. MLDv1 supports only any-source multicast (ASM). The implementation of source filter in IPv6 is similar to that in IPv4.

Charging Mode

For multicast services, carries or Internet service providers (ISPs) usually use two charging modes:

- Fixed charging. In this mode, programs are divided into different packages. Users need to pay a fixed amount of fee for each package in a fixed period (such as on a year or month basis). This charging mode does not restrict multicast users in the program-ordering count or the volume of traffic ordered.
- Pay per view (PPV): In this mode, users are charged according to how many times they order programs.

IPv6 multicast supports rights profiles and rights control and therefore supports fixed charging. IPv6 multicast does not support PPV because it does not support call detail record (CDR) reporting.

Program Ordering Behavior Analysis

IPv6 multicast supports local log recording and does not support log reporting. Log information can be obtained by local queries for analyzing program ordering behavior.

18.5.3 Differences Between IPv6 and IPv4 Multicast Features

IPv4 multicast can share VLANs with IPv6 multicast. Therefore, you can deploy IPv6 multicast in existing IPv4 multicast VLANs (MVLANs) just by enabling IPv6 multicast in the IPv4 MVLANs and adding IPv6 multicast programs to the rights profiles. The parameters that have already been configured for users, such as the bound rights profiles and MVLANs, remain unchanged.

IPv4 and IPv6 multicast programs can be added to the same MVLAN at the same time.

Table 18-17 lists the differences between IPv6 and IPv4 multicast features.

Table 18-17 Differences between IPv6 and IPv4 multicast features

Feature	IPv4	IPv6	Remarks
Protocol Independent Multicast-Source-Specific Multicast (PIM-SSM)	√	x	Only the OLT supports
VLAN-based multicast (TR101 multicast)	√	√	None
Multicast protocol	• Supports IGMPv1, IGMPv2, and IGMPv3, but does not support IGMPv3 Exclude mode. • Supports IGMP proxy. • Supports IGMP snooping, including	• Supports MLDv1 and MLDv2 but does not support MLDv2 Exclude mode. • Supports MLD proxy. • Supports MLD snooping, including	MLD snooping/proxy has similar functions as IGMP snooping/proxy and they are defined by RFC4541 and RFC4605. **NOTE** MLD does not support CTC

Feature	IPv4	IPv6	Remarks
	snooping with proxy.	snooping with proxy.	multicast.
Multicast preview	√	x	None
Multicast bandwidth call admission control (CAC)	√	√	None
Multicast call detail record (CDR)	√	x	None
MLD packet statistics measurement	√	√	None
Multicast program traffic query	√	√	None
Global leave	√	x	None
Load sharing between active and standby control boards	√	√	Only the OLT supports
Transparent transmission of multicast protocol packets (including support for transparent transmission policies, and processing policies in the event of a mismatch)	√	x	None
Remote acceptance for the multicast services provisioned through xDSL and OPFA boards	√	x	Only the OLT supports
Double VLAN tags for multicast packets transmitted upstream	√	√	Only the OLT supports
Aggregation and protection on multicast upstream ports and multicast cascade ports	√	√	Only the OLT supports
S+G forwarding	√	x	Only the OLT supports
Configuration of multicast query parameters on a VLAN basis	√	√	None
Local controllable multicast on the optical network terminal (ONT)	√	x	Only the ONT supports

Feature	IPv4	IPv6	Remarks
Dynamic generation of multiple program segments	√	x	None
Unicast GEM ports	√	√	Only the OLT supports

NOTE
In the preceding table, a tick (√) indicates that the feature is supported while a cross (x) indicates not supported.

18.6 Network Application

Multicast applies to FTTB/C, FTTH, and FTTO scenarios and ports connecting to multicast users.

Figure 18-52 GPON FTTx multicast network application

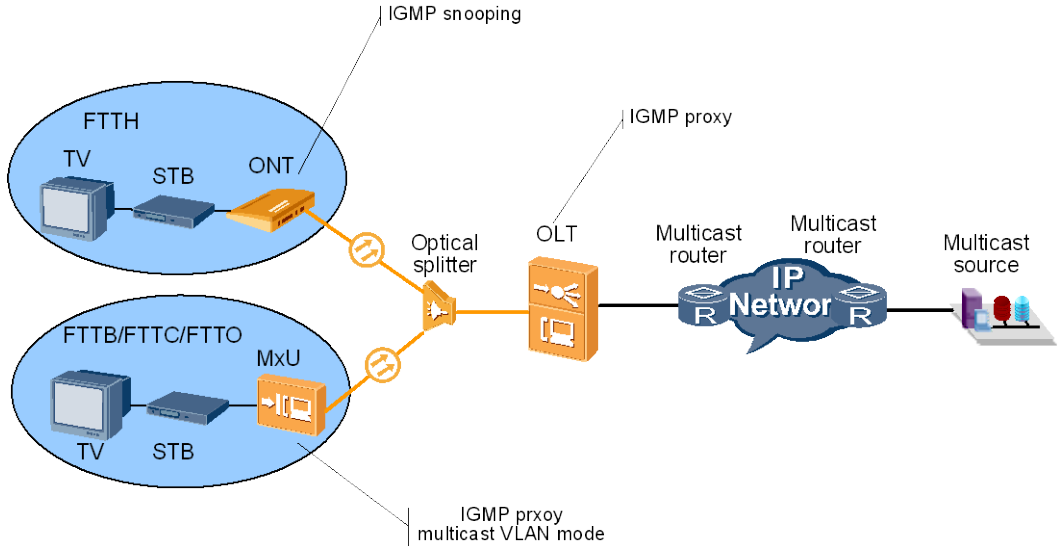
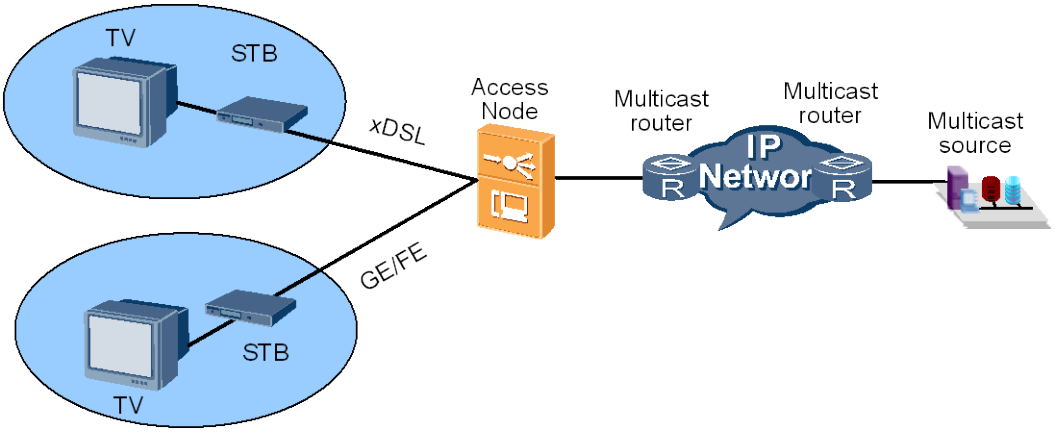


Figure 18-53 Network application for ports connecting to multicast users



18.7 Configuring the Multicast Service

The device supports multicast cascading for reducing the number of ports used on the convergence device, and also supports MSTP network protection. With these two functions, the network structure is optimized and the multicast service reliability is improved.

Context

To ensure service quality, physical layer retransmission is recommended. This retransmission technology is put forward to more reliably transmit services over lines. Compared with traditional data and voice services, video services such as IPTV and video on demand (VoD) impose far higher requirements on bit error ratio and packet loss ratio but lower requirements on delay.

- Run the **adsl extline-profile add(rfc2662)** or **adsl extline-profile modify(rfc2662)** command to configure physical layer retransmission for asymmetric digital subscriber line (ADSL) lines.
- Run the **vdsl channel-profile add(tr129)**, **vdsl channel-profile modify(tr129)**, **vdsl channel-profile quickadd(tr129)**, or **vdsl channel-profile quickmodify(tr129)** command to configure physical layer retransmission for very high speed digital subscriber line (VDSL) lines

18.7.1 Differences Between IPv4 and IPv6 Multicast Configurations

IPv6 multicast refers to the multicast service implemented using the IPv6 protocol. Differences between IPv4 and IPv6 multicast configurations mainly relate to commands and functions. Before configuring IPv6 multicast, it is recommended that you familiarize yourself with the procedures and principles of configuring IPv4 multicast and be aware of the differences between IPv4 and IPv6 multicast configurations.

Differences Regarding Commands

Table 18-18 Differences between IPv4 and IPv6 multicast commands

NOTE
The hyphen (-) indicates that the IPv4 command does not have a counterpart in IPv6. In other words, the corresponding multicast feature is not supported in IPv6.

IPv4	IPv6	Remark
Multicast additional functions		
igmp leave-proxy	igmp ipv6 leave-proxy	-
igmp priority	igmp ipv6 priority	-
imgp report-proxy	igmp ipv6 report-proxy	-
igmp accelerator	-	Only OLT supports
igmp echo	-	-
igmp encapsulation	-	-
igmp multicast-tag	igmp multicast-tag	Only OLT supports
igmp policy	-	-
igmp query-offline-user	-	-
igmp user-action-report	-	-
igmp mismatch	-	-
display igmp policy	-	-
igmp sip-gip-forward	-	Only OLT supports
Protocol parameters		
igmp proxy router gen-query-interval	igmp ipv6 router gen-query-interval	-
igmp proxy router gen-response-time	igmp ipv6 router gen-response-time	-
igmp proxy router robustness	igmp ipv6 router robustness	-
igmp proxy router sp-query-interval	igmp ipv6 router sp-query-interval	-
igmp proxy router sp-query-number	igmp ipv6 router sp-query-number	-
igmp proxy router sp-response-time	igmp ipv6 router sp-response-time	-
igmp initial-unsolicited-report interval	igmp ipv6 initial-unsolicited-report interval	-

IPv4	IPv6	Remark
igmp unsolicited-report interval	igmp ipv6 unsolicited-report interval	-
display igmp config vlan	display igmp ipv6 config vlan	-
igmp proxy router timeout	-	-
Multicast VLAN (MVLAN)		
igmp mode	igmp ipv6 mode	-
igmp match mode	igmp ipv6 match mode	-
igmp version	igmp ipv6 version	-
igmp match group	igmp ipv6 match group	-
display igmp config vlan	display igmp ipv6 config vlan	-
igmp inner-vlan	igmp inner-vlan	Only OLT supports
igmp send global-leave	-	-
Multicast users and rights		
debugging igmp	debugging igmp ipv6	-
Multicast preview		
igmp preview	-	-
igmp preview auto-reset-time	-	-
igmp preview reset count	-	-
igmp preview reset record	-	-
igmp preview-profile add	-	-
igmp preview-profile delete	-	-
igmp preview-profile modify	-	-
display igmp preview user	-	-
display igmp preview-profile	-	-
Statistics measurement		
igmp statistic reset	igmp ipv6 statistic reset	-
display igmp statistic	display igmp ipv6 statistic	-
display multicast flow-statistic	display multicast flow-statistic	-

IPv4	IPv6	Remark
Logs		
display igmp log statistic	-	Only OLT supports
display igmp log	-	-

NOTE
The multicast commands that are not listed in the table above are the commands supported by both IPv4 and IPv6.

Differences Regarding Functions

Compared with the IPv4 multicast feature, the IPv6 multicast feature does not support the following sub-features for OLT:

- SIP+GIP forwarding mode
- Layer 3 IPv6 multicast routing protocol
- Transparent transmission of unknown IGMP packets
- Spanish multicast log mode
- Dynamic generation of multiple program segments
- Multicast preview
- Video stream statistics measurement
- Global leave
- Load sharing between active and standby control boards
- Transparent transmission policy for packets, and VLAN-based forwarding policy for unmatched IGMP packets
- Identification of PPPoE-encapsulated MLD packets (only transparent transmission is supported for PPPoE-encapsulated MLD packets)
- Reporting of IPv6 multicast log

Compared with the IPv4 multicast feature, the IPv6 multicast feature does not support the following sub-features for ONU:

- Transparent transmission of unknown IGMP packets
- Video stream statistics measurement
- Global leave
- Transparent transmission policy for packets, and VLAN-based forwarding policy for unmatched IGMP packets
- Identification of PPPoE-encapsulated MLD packets (only transparent transmission is supported for PPPoE-encapsulated MLD packets)
- Reporting of IPv6 multicast log

Differences Regarding Multicast Basic Service Configurations

IPv4 multicast can share VLANs with IPv6 multicast. Therefore, you can deploy IPv6 multicast in existing IPv4 MVLANs just by enabling IPv6 multicast in the IPv4 MVLANs

and adding IPv6 multicast programs to the rights profiles. The parameters that have already been configured for users, such as the bound rights profiles and MVLANs, remain unchanged.

18.7.2 Configuring the Multicast Service on a Single NE

When the network structure is simple, the configuration of a single NE can meet multicast service requirements. Compared with cascading networking, single-NE networking is more secure and stable, and provides more bandwidth resources, but requires more line resources. The method of configuring multicast services for an NE in the cascading or MSTP networking scenario is the same as that in single-NE networking scenario.

The default configuration of multicast is as follows:

Table 18-19 Default settings of the multicast service

Feature	Default Settings
Multicast protocol	Disabled
IGMP version	V3
Multicast program configuration mode	Static configuration mode
Multicast bandwidth management	Enabled
Multicast preview	Enabled
Multicast log function	Enabled

Prerequisite

The upper-layer device and multicast source have been configured.

Configuration Flowchart

Configuring Multicast Global Parameters

This topic describes the configuration of L2 multicast protocols (including IGMP proxy and IGMP snooping). For IPv4, the MA5621 supports the global configuration and the configuration based on the MVLAN. For IPv6, the MA5621 only supports the configuration based on the MVLAN.

Context

The multicast global parameters include general query, group-specific query, the policy of processing multicast packets and the multicast forwarding mode.

The description of a general query is as follows:

- Purpose: A general query packet is periodically sent by the access device to check whether there is any multicast user who leaves the multicast group without sending the leave packet. Based on the query result, the access device periodically updates the multicast forwarding table and releases the bandwidth of the multicast user that has left the multicast group.

- Principle: The access device periodically sends the general query packet to all online IGMP users. If the access device does not receive the response packet from a multicast user within a specified time (Robustness variable x General query interval + Maximum response time of a general query), it regards the user as having left the multicast group and deletes the user from the multicast group.

The description of a group-specific query is as follows:

- Purpose: A group-specific query packet is sent by the access device after a multicast user that is not configured with the quick leave attribute sends the leave packet. The group-specific query packet is used to check whether the multicast user has left the multicast group.
- Principle: When a multicast user leaves a multicast group, for example, switches to another channel, the user unsolicitedly sends a leave packet to the access device. If the multicast user is not configured with the quick leave attribute, the access device sends a group-specific query packet to the multicast group. If the access device does not receive the response packet from the multicast user within a specified duration (Robustness variable x Group-specific query interval + Maximum response time of a group-specific query), it deletes the multicast user from the multicast group.

NOTE

The configuration steps for IPv4 multicast and IPv6 multicast are similar but detailed commands are different. This topic describes the configuration steps for IPv4 multicast, and provides the configuration example for IPv6 multicast.

Table 18-20 lists the default settings of the multicast global parameters. In the actual application, you can modify the values according to the data plan.

Table 18-20 Default settings of the multicast global parameters

Parameter	Default Value
General query parameter	Query interval: 125s Maximum response time: 10s Robustness variable (query times): 2
Group-specific query parameter	Query interval: 1s Maximum response time: 0.8s. Robustness variable (query times): 2
Policy of processing multicast packets	IGMP packet: normal (IGMP packets are processed as controllable multicast) Unknown multicast packet: discard
Policy of processing multicast packets	IGMP packet: normal (IGMP packets are processed as controllable multicast) Unknown multicast packet: discard
Policy of processing multicast packets	IGMP packet: normal (IGMP packets are processed as controllable multicast) Unknown multicast packet: • For switch-oriented traffic streams: discard • For connection-oriented traffic streams: transparent transmission

Parameter	Default Value
Multicast forwarding mode (Only OLT supports)	disable(VLAN+GMAC mode)

Procedure

Configure the general query parameters.

1. Run the **igmp proxy router gen-query-interval** command to set the general query interval. By default, the general query interval is 125s.
2. Run the **igmp proxy router gen-response-time** command to set the maximum response time of the general query. By default, the maximum response time of the general query is 10s.
3. Run the **igmp proxy router robustness** command to set the robustness variable (query times) of the general query. By default, the robustness variable (query times) is 2.

Step 1 Set the group-specific query parameters.

1. Run the **igmp proxy router sp-response-time** command to set the group-specific query interval. By default, the group-specific query interval is 1s.
2. Run the **igmp proxy router sp-query-interval** command to set the maximum response time of the group-specific query. By default, the maximum response time of the group-specific query is 0.8s.
3. Run the **igmp proxy router sp-query-number** command to set the robustness variable (query times) of the group-specific query. By default, the robustness variable (query times) is 2.

Step 2 Configure the policy of processing multicast packets.

By default, the normal mode for processing IGMP packets is adopted. In this mode, IGMP packets are processed as controllable multicast. The discard mode is adopted for unknown multicast packets. In this mode, unknown multicast packets are discarded.

The default values are adopted for multicast service and do not need to be modified. To control the forwarding of multicast packets when configuring other services, run the following commands to configure the policy.

1. Run the **igmp policy** command to set the policy of processing IGMP packets.
2. Run the **multicast-unknown policy** command to set the policy of processing unknown multicast packets(downstream UDP packets).

Step 3 (Optional) Configure the multicast forwarding mode. (Only OLT supports)

The multicast forwarding mode including these following two types. Run the **igmp sip-gip-forward** { **disable** | **enable** } command to set the multicast forwarding mode.

- disable: Sets the forwarding mode to VLAN+GMAC. That is, packets are forwarded based on the MAC address mapped from the IP address of the multicast program and the multicast VLAN ID.
- enable: Sets the forwarding mode to SIP+GIP. That is, IGMP packets are forwarded in the mode of multicast source multicst VLAN ID+IP address+multicast program IP address. When multiple multicast sources have the programs with the same IP address,

to differentiate programs and their corresponding multicast sources, you need to set the forwarding mode of IGMP packets to SIP+GIP.

Step 4 Run the **display igmp config global** command to check whether the values of the multicast parameters are correct.

----End

Example

(IPv4) To configure the multicast general query parameters by setting the query interval to 150s, maximum response time to 20s, and number of queries to 3 on the multicast VLAN 100, do as follows:

```
huawei(config)#multicast-vlan 100
huawei(config-mvlan100)#igmp proxy router gen-query-interval 150
huawei(config-mvlan100)#igmp proxy router gen-response-time v3 20
huawei(config-mvlan100)#igmp proxy router robustness 3
```

(IPv4) To configure the multicast group-specific query parameters by setting the query interval to 200s, maximum response time to 100s, and number of queries to 3 on the multicast VLAN 100, do as follows:

```
huawei(config)#multicast-vlan 100
huawei(config-mvlan100)#igmp proxy router sp-query-interval 200
huawei(config-mvlan100)#igmp proxy router sp-response-time v3 100
huawei(config-mvlan100)#igmp proxy router sp-query-number 3
```

(IPv6) To configure the multicast general query parameters by setting the query interval to 150s, maximum response time to 20s, and number of queries to 3 on the multicast VLAN 200, do as follows:

```
huawei(config)#multicast-vlan 200
huawei(config-mvlan200)#igmp ipv6 router gen-query-interval 150
huawei(config-mvlan200)#igmp ipv6 router gen-response-time v2 20
huawei(config-mvlan200)#igmp ipv6 router robustness 3
```

(IPv6) To configure the multicast group-specific query parameters by setting the query interval to 200s, maximum response time to 100s, and number of queries to 3 on the multicast VLAN 200, do as follows:

```
huawei(config)#multicast-vlan 200
huawei(config-mvlan200)#igmp ipv6 router sp-query-interval 200
huawei(config-mvlan200)#igmp ipv6 router sp-response-time v2 100
huawei(config-mvlan200)#igmp ipv6 router sp-query-number 3
```

Configuring the Multicast VLAN and the Multicast Program

In the application of multicast service, multicast VLANs (MVLANs) are used to distinguish multicast ISPs. Generally, an MVLAN is allocated to each multicast ISP for the VLAN-based management of multicast programs, multicast protocols, IGMP versions, and the VLAN-based control of multicast domain and user right.

Context

To create a multicast VLAN, a common VLAN must be created first. The multicast VLAN can be the same as the unicast VLAN. In this case, the two VLANs can share the same service

stream channel. The multicast VLAN can be different from the unicast VLAN. In this case, the two VLANs use different service stream channels.

One user port can be added to multiple multicast VLANs under the following restrictions:

- Among all the multicast VLANs of a user port, only one multicast VLAN is allowed to have dynamically generated programs.
- One user port is not allowed to belong to multiple MVLANs that are in the IGMP v3 snooping mode.

The source IP address in the multicast packets that are sent to the upper device by the access device may be as follows:

- If the IP address of the program VLAN interface is configured, the source IP address is the IP address of VLAN interface.
- If the IP address of the program VLAN interface is not configured, the source IP address is the host IP address of the program.
- If the host IP address is not configured, the default address 0.0.0.0 is used.

Table 18-21 lists the default settings of the MVLAN attributes, including the Layer 2 multicast protocol, IGMP version, multicast program, and multicast upstream port.

Table 18-21 Default settings of the MVLAN attributes

Parameter	Default Value
Program matching mode	enable (static configuration mode)
Multicast upstream port mode	default
Layer 2 multicast protocol	off (multicast function disabled)
IGMP version	v3
Priority of forwarding IGMP packets by the upstream port	6
Group filter mode (Only OLT supports)	asm-ssm

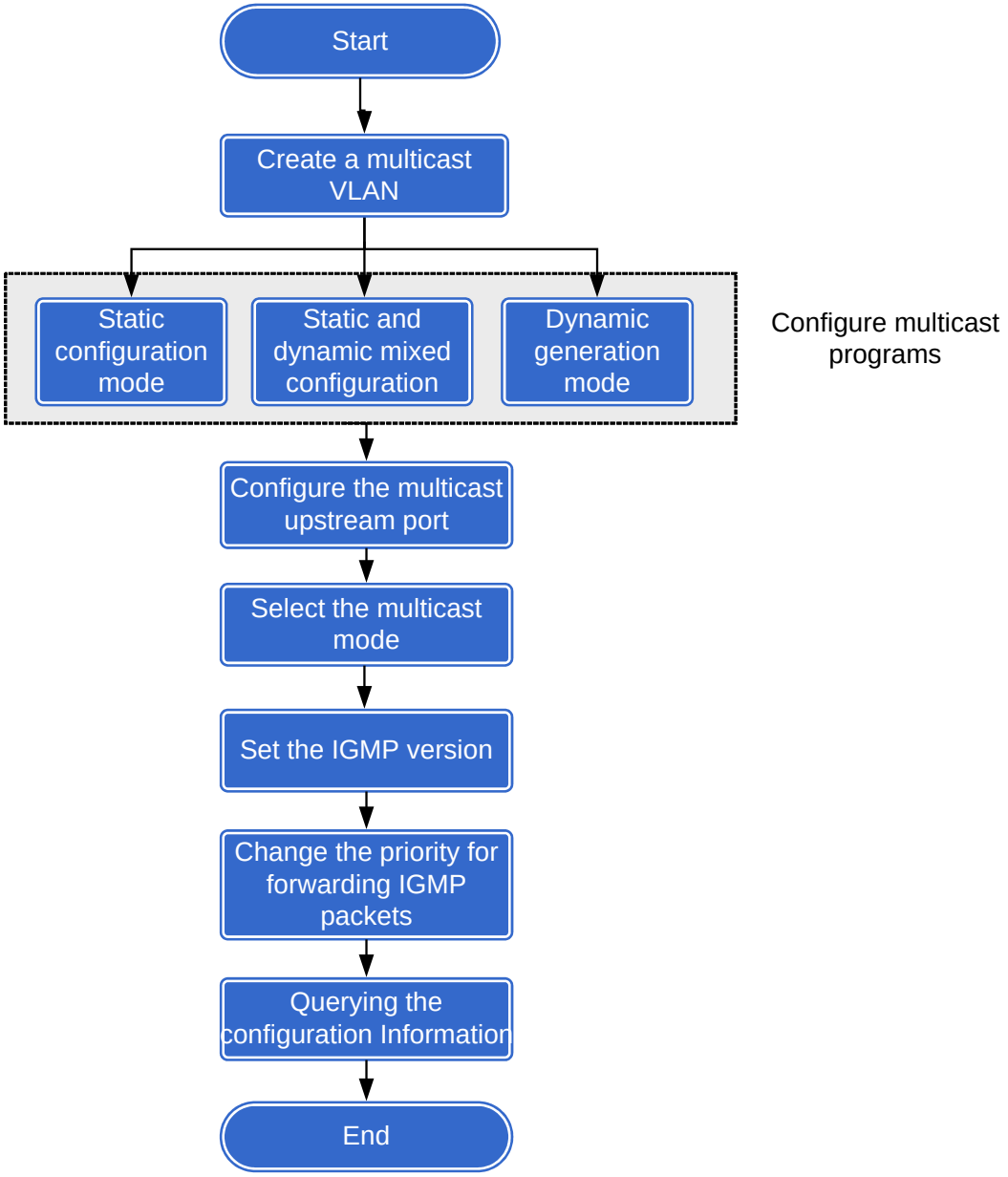
NOTE

The device supports IPv4 and IPv6 multicast services. The two types of services have the same VLAN configurations and only differ in commands. The procedure in this topic uses IPv4 multicast as an example, and the example uses IPv6 multicast as an example.

Configuration Flowchart

Figure 18-54 shows the flowchart for configuring the multicast VLAN and multicast program.

Figure 18-54 Flowchart for configuring the multicast VLAN and multicast program



Procedure

Create a multicast VLAN.

1. Run the **vlan** command to create a VLAN, and set the VLAN type according to the actual application.
2. Run the **multicast-vlan** command to set the created VLAN to a multicast VLAN. The VLAN with S+C forwarding mode cannot be set as a multicast VLAN.

Step 1 Configure multicast programs.

The multicast VLAN can be configured statically or generated dynamically. The program configuration of the MVLAN has three modes: static configuration, dynamic generation, and static and dynamic mixed configuration.

- Static configuration mode: Configure the program list before users watch the video programs. In this mode, the right profile can be used to implement controllable multicast. The program list and the right profile, however, need to be maintained according to the change of the video service. The program host, program prejoin, and multicast bandwidth management functions are supported.

1. Run the **igmp match mode enable** or **igmp ipv6 match mode enable** command to set the static configuration mode. By default, the system adopts the static configuration mode.
2. Run the **igmp program add** command to add a multicast program.

NOTE

If the IGMP version of a multicast VLAN is v3, the program must be configured with a source IP address. If the IGMP version of a multicast VLAN is v2, the program must not be configured with a source IP address.

3. Add a right profile.

   In the BTV mode, run the **igmp profile add** command to add a right profile.
4. Bind the program to the right profile.

   In the BTV mode, run the **igmp profile** command to bind the program to the right profile, and set the right to watch.

NOTE

When a user is bound to multiple right profiles, and the right profiles have different rights to a program, the right with the highest priority prevails. You can run the **igmp right-priority** command to adjust the priorities of the four rights: watch, preview, forbidden, and idle. By default, the priorities of the four rights are forbidden > preview > watch > idle.

- Dynamic generation mode: A program list is dynamically generated according to the programs requested by users. In this mode, the program list does not need to be configured or maintained; however, the functions such as program management, user multicast bandwidth management, program preview, and program prejoin are not supported.

1. Run the **igmp match mode disable** or **igmp ipv6 match mode disable** command to set the dynamic generation mode.

NOTICE

The **igmp match mode** command can be executed only when the IGMP mode is disabled.

2. Run the **igmp match group** or **igmp ipv6 match group** command to configure the IP address range of the program group that can be dynamically generated. Users can order only the programs whose IP addresses are within the specified range.

- Static and dynamic mixed configuration: Configure the program list before users watch the video programs. In this mode, the right profile can be used to implement controllable multicast. The program list and the right profile, however, need to be maintained according to the change of the video service. The program host, program prejoin, and multicast bandwidth management functions are supported.

1. Run the **igmp match mode disable** or **igmp ipv6 match mode disable** command to set the mode to the dynamic generation mode.

2. Run the **igmp match group** or **igmp ipv6 match group** command to configure the IP address range of the program group that can be dynamically generated. Users can order only the programs whose IP addresses are within the specified range.
3. Run the **igmp program add** [**name** *name*] **ip** *ip-addr* [**sourceip** *ip-addr*] [**hostip** *ip-addr*] command to add a multicast static program.

NOTE

When the range of static program IP addresses and the range of dynamic program IP addresses overlap each other, static programs can go online with priority.

4. Run the **igmp group-filter-mode** command to set the group filter mode based on multicast VLAN (MVLAN). (Only OLT supports)

NOTE

- When the group filter mode of an MVLAN is configured to **asm-only** or **asm-ssm**, only one program with the unique multicast IP address is generated in the MVLAN. The [*, G] multicast forwarding table is used for this MVLAN instance on the forwarding plane.
- When the group filter mode of an MLVAN is configured to **ssm-only**, multiple programs with the same multicast IP addresses but different source IP addresses can be generated in the MVLAN. The [s, g] multicast forwarding table is used for this VLAN instance on the forwarding plane.

NOTE

The source IP addresses are regarded as different ones when they have different least significant 20 bits from each other.

- The maximum number of programs is calculated according to the number of actually-generated programs. For example:
- When a multicast user joins an MVLAN with the multicast filter mode **asm-ssm** and the system receives two packets with IP addresses [S1, G1] and [S2, G1], the system generates only one multicast program with the multicast IP address G1 for the multicast user;
- When a multicast user joins an MVLAN with the multicast filter mode **ssm-only** and the system receives two packets with IP addresses [S1, G1] and [S2, G1], the system generates two multicast programs with IP addresses [S1, G1] and [S2, G1].

Step 2 Configure the multicast upstream port.

1. Run the **igmp uplink-port** command to configure the multicast upstream port. The packets of the MVLAN corresponding to the upstream port are forwarded and received by this upstream port.
2. In the BTV mode, run the **igmp uplink-port-mode** command to change the mode of the multicast upstream port. By default, the port is in the default mode. In the MSTP network, the port adopts the MSTP mode.
   - Default mode: If the MVLAN contains only one upstream port, the multicast packets that go upstream can be sent only by this port. If the MVLAN contains multiple upstream ports, the multicast packets that go upstream are sent by all the upstream ports.
   - MSTP mode: This mode is adopted in the MSTP network.

Step 3 Select the multicast mode.

Run the **igmp mode** { **proxy** | **snooping** } command to select the Layer 2 multicast mode. By default, the multicast mode is disabled.

In terms of multicast processing mode, the access device supports the Internet Group Management Protocol (IGMP) Proxy and IGMP Snooping Layer 2 multicast protocols. IGMP proxy and IGMP snooping both support multicast video data forwarding; however, the two modes have different processing mechanisms.

- In IGMP snooping, the related information for maintaining multicast forwarding entries is obtained by listening to the IGMP packets between the user and the multicast router.

- IGMP proxy intercepts the IGMP packets between the user and the multicast router, processes the IGMP packets, and then forwards the IGMP packets to the upper-layer multicast router. For the multicast user, the access device is a multicast router that implements the router functions in the IGMP protocol; for the multicast router, the access device is a multicast user.

In the IGMP snooping mode, proxy can be enabled for the report packet and the leave packet. When a multicast user joins or leaves a multicast program, the access device can implement IGMP proxy. IGMP snooping and IGMP proxy are controlled separately.

- Run the **igmp report-proxy enable** command to enable the proxy of the snooping report packet. When the first user requests to join a program, after authenticating the user, the access device sends the user report packet to the network side and receives a corresponding multicast stream from the multicast router. The report packets of the users that follow the first user are not sent by the access device to the network side.
- Run the **igmp leave-proxy enable** command to enable the proxy of the snooping leave packet. When the last user requests to leave the program, the access device sends the user leave packet to the network side to request the upper-layer device to stop sending multicast streams. The leave packets of the users that precede the last user are not sent by the access device to the network side.

Step 4 Set the IGMP version.

Run the **igmp version**{ **v2** | **v3** } command to set the IGMP version. By default, IGMP v3 is enabled in the system. If the upper-layer and lower-layer devices in the network are IGMP v2 devices and cannot recognize the IGMP v3 packets, run this command to change the IGMP version.

NOTICE

When you run the **igmp version**{ **v2** | **v3** } command to configure the IGMP version:

- This command can be executed only when the IGMP mode is disabled.
- The system will delete the programs with source IP addresses in this multicast VLAN. In this case, if a user is online, the system will force the user to go offline.

IGMP v3 is compatible with IGMP v2 in packet processing. If IGMP v3 is enabled on the access device and the upper-layer multicast router switches to IGMP v2, the access device automatically switches to IGMP v2 when receiving the IGMP v2 packets. If the access device does not receive any more IGMP v2 packets within the preset IGMP v2 timeout time, it automatically switches back to IGMP v3. In the BTV mode, run the **igmp proxy router timeout** command to set the IGMP v2 timeout time. By default, the timeout time is 400s.

Step 5 Change the priority for forwarding IGMP packets.

Run the **igmp priority** command to change the priority for forwarding the IGMP packets by the upstream port. By default, the priority is 6 and does not need to be changed.

- In the IGMP proxy mode, the IGMP packets sent from the upstream port to the network side adopt the priority set through the preceding command in the MVLAN.
- In the IGMP snooping mode, the IGMP packets forwarded to the network side adopt the priority of the user service stream. The priority of the service stream is set through the traffic profile.

Step 6 Check whether the configuration is correct.

- Run the **display igmp config vlan** command to query the attributes of the multicast VLAN.
- Run the **display igmp program vlan** command to query the information about the program of the MVLAN.

----End

Example

Assume that:

- MVLAN ID: 101
- Program configuration mode: static configuration; program IP address: 224.1.1.1
- Source IP address: 10.10.10.10; host IP address: 10.0.0.254
- Program bandwidth: 5000 kbit/s
- MVLAN upstream port: 0/19/0
- Protocol: IGMP proxy; IGMP version: v3
- Group filter mode: ssm-only (Only OLT supports)

To configure the MVLAN and multicast program for the IPv4 multicast, do as follows:

```
huawei(config)#vlan 101 smart
huawei(config)#multicast-vlan 101
huawei(config-mvlan101)#igmp match mode enable
huawei(config-mvlan101)#igmp program add name movie ip 224.1.1.1 sourceip 10.10.10.10
hostip 10.0.0.254 bandwidth 5000
huawei(config-mvlan101)#igmp uplink-port 0/19/0
huawei(config-mvlan101)#igmp mode proxy
  Are you sure to change IGMP mode?(y/n)[n]:y
  Command is being executed. Please wait...
  Command has been executed successfully
huawei(config-mvlan101)#igmp version v3
huawei(config-mvlan101)#igmp group-filter-mode ssm-only
```

Assume that:

- MVLAN ID: 101
- Program configuration mode: dynamic generation
- Address range of the dynamic program group: 224.1.1.10 to 224.1.1.50
- Program bandwidth: 5000 kbit/s
- MVLAN upstream port: 0/19/0
- Protocol: IGMP proxy; IGMP version: v3

To configure the MVLAN and multicast program for the IPv4 multicast, do as follows:

```
huawei(config)#vlan 101 smart
huawei(config)#multicast-vlan 101
huawei(config-mvlan101)#igmp match mode disable
  This operation will delete all the programs in current multicast vlan
  Are you sure to change current match mode? (y/n)[n]: y
  Command is being executed, please wait...
  Command has been executed successfully
huawei(config-mvlan101)#igmp match group ip 224.1.1.10 to-ip 224.1.1.50
```

```
huawei(config-mvlan101)#igmp uplink-port 0/19/0
huawei(config-mvlan101)#igmp mode proxy
  Are you sure to change IGMP mode?(y/n)[n]:y
  Command is being executed. Please wait...
  Command has been executed successfully
huawei(config-mvlan101)#igmp version v3
```

Assume that:

- MVLAN ID: 101
- Program configuration mode: static and dynamic mixed configuration
- MVLAN upstream port: 0/19/0
- IP address of the static program: 224.1.1.1; source IP address: 10.10.10.10; host IP address: 10.0.0.254; program bandwidth: 5000 kbit/s
- Address range of the dynamic program group: 224.1.1.10 to 224.1.1.50
- Protocol: IGMP proxy; IGMP version: v3

To configure the MVLAN and multicast program for the IPv4 multicast, do as follows:

```
huawei(config)#vlan 101 smart
huawei(config)#multicast-vlan 101
huawei(config-mvlan101)#igmp match mode disable
  This operation will delete all the programs in current multicast vlan
  Are you sure to change current match mode? (y/n)[n]: y
  Command is being executed, please wait...
  Command has been executed successfully
huawei(config-mvlan101)#igmp match group ip 224.1.1.10 to-ip 224.1.1.50
huawei(config-mvlan101)#igmp program add name movie ip 224.1.1.1 sourceip 10.10.10.10
hostip 10.0.0.254 bandwidth 5000
huawei(config-mvlan101)#igmp uplink-port 0/19/0
huawei(config-mvlan101)#igmp mode proxy
  Are you sure to change IGMP mode?(y/n)[n]:y
  Command is being executed. Please wait...
  Command has been executed successfully
huawei(config-mvlan101)#igmp version v3
```

Assume that:

- MVLAN ID: 101
- Program configuration mode: static configuration; program IPv6 address:ffff::1
- Source IPv6 address: 2000::1
- Program bandwidth: 5000 kbit/s
- MVLAN upstream port: 0/19/0
- Protocol: IGMP proxy; IGMP version: v2

To configure the MVLAN and multicast program for the IPv6 multicast, do as follows:

```
huawei(config)#vlan 101 smart
huawei(config)#multicast-vlan 101
huawei(config-mvlan101)#igmp ipv6 match mode enable
huawei(config-mvlan101)#igmp program add name movie ipv6 ffff::1 source-ipv6 2000::1
 bandwidth 5000
huawei(config-mvlan101)#igmp uplink-port 0/19/0
```

```
huawei(config-mvlan101)#igmp ipv6 mode proxy
  Are you sure to change IGMP mode?(y/n)[n]:y
  Command is being executed. Please wait...
  Command has been executed successfully
huawei(config-mvlan101)#igmp ipv6 version v2
```

Configuring the Multicast GPON ONT

When the MA5600T/MA5603T/MA5608T is connected with an ONT or an MDU, you need to configure the multicast interconnection data to forward the multicast traffic streams.

Prerequisites

Before configuring the multicast GPON ONT, you must add the ONT correctly. For the configuration method, see 2.13.3 Configuring a GPON ONT (Profile Mode).

Context

- When the OLT is connected with an ONT such as the HG850e, the MA5600T/MA5603T/MA5608T manages the ONT in the OMCI mode. In this case, you need to configure the ONT line profile and the ONT service profile, configure the multicast data in the ONT service profile, and bind the profiles to the ONT to issue the multicast service.
- When the OLT is connected with an MDU such as the MA5620 or MA5616, the MA5600T/MA5603T/MA5608T manages the MDU in the SNMP mode. In this case, you do not need to configure the ONT service profile. You only need to configure the multicast data on the MDU interconnected with the MA5600T/MA5603T/MA5608T to forward the multicast traffic streams.

Procedure

Add an ONT line profile.

For the configuration method, see Configuring a GPON ONT Line Profile (Profile Mode).

Step 1 Add an ONT service profile.

Run the **ont-srvprofile gpon** command to add a GPON ONT service profile, and then enter the GPON ONT service profile mode.

If the ONT management mode is the SNMP mode, you do not need to configure the service profile. After adding a GPON ONT service profile, directly enter the GPON ONT service profile mode to configure the related multicast data.

1. Run the **ont-port** command to configure the port capability set of the ONT. The port capability set in the ONT service profile must be the same as the actual ONT capability set.
2. Run the **port vlan** command to configure the port VLAN of the ONT.
3. Configure the multicast mode of the ONT.

   Run the **multicast mode** { **igmp-snooping**|**olt-control**|**unconcern** } command to select the multicast mode.

   - **igmp-snooping**: IGMP snooping obtains related information and maintains the multicast forwarding entries by listening to the IGMP packets in the communication between the user and the multicast router.

- **olt-control**: It is the dynamic controllable multicast mode. A multicast forwarding entry can be created for the multicast join packet of the user only after the packet passes the authentication.
- **unconcern**: It is the unconcern mode. After this mode is selected, the OLT does not limit the multicast mode, and the multicast mode on the OLT automatically matches the multicast mode on the ONT.

4. (Optional)Configure the multicast forwarding mode.

   Run the **multicast-forward** { **untag** | **tag** { **translation***vlanid* | **transparent** } | **unconcern** command to configure the multicast forwarding mode and multicast forwarding VLAN. The forwarding mode is not concerned by default.

   - **tag**: Specifies the multicast forwarding mode as tag. If the VLAN tag of the multicast packet needs to be transparently transmitted, use **transparent**; if the VLAN tag of the multicast packet needs to be switched, use **translation** and set the VLAN tag used after the switching. When the ONT is directly connected to the home gateway in the application, use this parameter.
   - **untag**: Specifies the multicast forwarding mode as untag, that is, the downstream multicast packet from the ONT's Ethernet port to a next directly connected device does not carry the VLAN tag. When the ONT is directly connected to the set top box (STB) or PC, use this parameter.
   - **unconcern**: Indicates that the multicast forwarding mode is not concerned. When the ONT multicast mode need not be configured by the OLT and is determined by the ONT condition, use **unconcern**. This value is the default value.

5. After the configuration is complete, run the **commit** command to make the configured service profile take effect.

NOTE

For an ONT that is added through the **ont add** command or an automatically found ONT that is confirmed through the **ont confirm** command, if you run the **commit** command after modifying the ONT line profile parameters and the ONT service profile parameters, the modified profile parameters take effect immediately.

----End

Example

To configure the ONT service profile 10 of 4 ETH ports, 2 POTS ports, the VLAN of the ETH port as 10, the multicast mode as IGMP snooping, the multicast forwarding mode as unconcern, do as follows:

```
huawei(config)#ont-srvprofile gpon profile-id 10
huawei(config-gpon-srvprofile-10)#ont-port eth 4 pots 2
huawei(config-gpon-srvprofile-10)#port vlan eth 1 10
huawei(config-gpon-srvprofile-10)#multicast mode igmp-snooping
huawei(config-gpon-srvprofile-10)#multicast-forward unconcern
huawei(config-gpon-srvprofile-10)#commit
huawei(config-gpon-srvprofile-10)#quit
```

Configuring a Multicast User

This topic describes how to configure a multicast user and the related user right for provisioning the multicast service.

Prerequisites

Before configuring a multicast user, create a service channel. The procedure is as follows:

1. 13.3.9 Configuring a VLAN
2. Configuring an Upstream Port
3. Configuring ADSL2+ User Ports or Configuring VDSL2 User Ports
4. Creating an xDSL or Ethernet Service Flow

NOTE
The multicast service supports the IPoE and PPPoE user access modes, but does not support the IPoEoA or PPPoEoA user access mode.

- Configure a GPON multicast user
  a. 13.3.9 Configuring a VLAN
  b. Configuring an Upstream Port
  c. Configuring the Multicast GPON ONT
  d. 2.13.4 Configuring a GPON Port
  e. Creating a GPON Service Flow (in Profile Mode with Universal Configurations)

Context

Add a multicast user, and bind the multicast user to the multicast VLAN to create a multicast member. Bind the multicast user to a right profile to implement multicast user authentication.

Table 18-22 lists the default settings of the attributes related to the multicast user.

Table 18-22 Default settings of the attributes related to the multicast user

Parameter	Default Value
Maximum number of programs that can be watched by the multicast user	8
Maximum number of programs of different priorities that can be watched by the multicast user	no-limit
Quick leave mode of the multicast user	MAC-based
Global switch of multicast user authentication	enable
IGMP version of the multicast user	v3

Procedure

In the global config mode, run the **btv** command to enter the BTV mode.

Step 1 Configure a multicast user and the multicast user attributes.

1. Add a multicast user.

   Run the **igmp user add service-port** command to add a multicast user.

2. Configure the maximum number of programs that can be watched by the multicast user.
   - Run the **igmp user add service-port** *index* **max-program** { **max-program-num** | **no-limit** } command to configure the maximum number of programs that can be watched by the multicast user concurrently.
   - Run the **igmp user watch-limit service-port** { **hdtv** | **sdtv** | **streaming-video** } command to configure the maximum number of programs of different priorities that can be watched by the multicast user.

3. Set the quick leave mode of the multicast user.

   Run the **igmp user add** or the **igmp user modify** command with the **quickleave { immediate | disable | mac-based }** parameter to configure the leave mode of the multicast user. By default, the leave mode is the mac-based mode.
   - **Immediate**: After receiving the leave packet of the multicast user, the system immediately deletes the multicast user from the multicast group.
   - **Disable**: After receiving the leave packet of the multicast user, the system sends an ACK packet to confirm that the multicast user leaves, and then deletes the multicast user from the multicast group.
   - **MAC-based**: It is the quick leave mode based on the MAC address. The system checks the MAC address in the leave packet of the user. If it is the same as the MAC address in the report packet of the user and it is the last MAC addres of multicast user, the system immediately deletes the multicast user from the multicast group. Otherwise, the system does not delete the multicast user. This mode is applied to the scenario with multiple terminals.

4. Set the IGMP version for the multicast user.

   Run **igmp user add service-port** *index* **igmp-version** { **v2** | **v3** | **v3-forced** | **v2-with-query** } or **igmp user add port***frameid/slotid/portid***igmp-version** { **v2** | **v3** | **v3-forced** | **v2-with-query** } command to set the IGMP version for the multicast user. Each multicast users has an independent querier instance. This command specifies the IGMP version (default: v3) for the multicast user querier.
   - v2: specifies the IGMP version to v2 for the multicast user querier. When this setting applies, the system processes only IGMP v2 packets and directly drops IGMP v1 packets and IGMP v3 packets.
   - v3: specifies the v3–compatible mode (default setting for the system). When this setting applies, the system automatically specifies the IGMP version according to the version of the IGMP packets sent by users, but it directly drops IGMP v1 packets.
   - v3-forced: forcibly specifies the IGMP version to V3 for the multicast user querier. When this setting applies, the system processes only IGMP v3 packets but directly drops IGMP v1 packets and IGMP v2 packets.
   - v2-with-query: specifies the user querier version to IGMP v2 and can receive the IGMP v3 packet sent from the user side. When the system receives an IGMP v3 packet sent from the user side, the system sends a common query packet of the IGMP v2 version and discards the IGMP v3 packet.

Step 2 Configure multicast user authentication.

To control the right of a multicast user, you can enable the multicast user authentication function.

1. Configure the multicast user authentication function.

Run the **igmp user add** or the **igmp user modify** command with the **{ auth | no-auth }** parameter to configure whether to authenticate a multicast user.

NOTE

After configuring multicast user authentication, you need to enable the global authentication function to make the configuration take effect. By default, the global authentication function of multicast user is enabled. You can run the **igmp proxy authorization** command to change the configuration.

2. Bind the multicast user to the right profile. This operation is to implement user authentication.

   Run the **igmp user bind-profile** command to bind the user to a right profile. After the binding, the multicast user has the rights to the programs as configured in the profile.

Step 3 Bind the multicast user to a multicast VLAN.

In the multicast VLAN mode, run the **igmp multicast-vlan member** command to bind the user to the multicast VLAN. Then, the multicast user becomes a multicast member of the multicast VLAN and can request the programs configured in the multicast VLAN.

Step 4 Run the **display igmp user** command to check whether the related multicast user information is correctly configured.

----End

Example

To add multicast user (port) 0/2/1 to multicast VLAN 101, enable user authentication, enable log report, set the maximum bandwidth to 10 Mbit/s, set IGMP version of the multicast user to v3-forced, and bind the user to right profile **music**, do as follows:

```
huawei(config)#service-port 100 vlan 101 adsl 0/2/1 vpi 0 vci 35 rx-cttr 2 tx-cttr 2
huawei(config)#btv
huawei(config-btv)#igmp user add service-port 100 auth log enable max-bandwidth 10240
igmp-version v3-forced
huawei(config-btv)#igmp user bind-profile service-port 100 profile-name music
huawei(config-btv)#quit
huawei(config)#multicast-vlan 101
huawei(config-mvlan10)#igmp multicast-vlan member service-port 100
```

(Optional) Configuring the Multicast Bandwidth

To limit the multicast bandwidth of a user, you can enable multicast bandwidth management, that is, connection admission control (CAC), and then control the bandwidth of a multicast user by setting the program bandwidth and the user bandwidth.

Prerequisites

The program matching mode of the multicast VLAN must be the static configuration mode.

Context

If the CAC function, not the dynamic Access Node Control Protocol (ANCP) CAC function is enabled, and a user requests for a multicast program, the system compares the remaining bandwidth of the user (bandwidth configured for the user – total bandwidth of the online programs of the user) with the bandwidth of the multicast program. The system determines whether the user can watch the multicast program based on the result:

- If the remaining bandwidth of the user is sufficient, the system adds the user to the multicast group.
- If the bandwidth is insufficient, the system does not respond to the request of the user. That is, the request of the user fails.

The MA5600T/MA5603T/MA5608T also supports dynamic application of unicast bandwidth for xDSL multicast users. This function is implemented through the ANCP protocol between the MA5600T/MA5603T/MA5608T and the BRAS. If CAC and ANCP are enabled on the device, when the user requests a multicast program, the resource admission control subsystem (RACS) compares the remaining bandwidth of the user with the bandwidth of the requested program. If the remaining bandwidth of the user is sufficient, the RACS adds the user to the multicast group. If the remaining bandwidth of the user is insufficient, the MA5600T/MA5603T/MA5608T applies to the RACS for the multicast video resource. Then, the RACS determines whether the available unicast video bandwidth of the user port can be allocated for the multicast video service of the user. In this way, the video bandwidth of the port is adjusted dynamically. For details on the ANCP configuration, see 24.7.3 Configuring ANCP.

If the CAC function is disabled, the system does not guarantee the bandwidth of the multicast program. When the bandwidth is not guaranteed, problems such as mosaic and delay occur in the multicast program.

Table 18-23 lists the default settings of the CAC parameters.

Table 18-23 Default settings of the CAC parameters

Parameter	Default Value
Global CAC function	enable
Bandwidth of the multicast program	5000 kbit/s
Bandwidth of the multicast user	no-limit
Bandwidth of the GPON port	716800 kbit/s

Procedure

In the global config mode, run the **btv** command to enter the BTV mode.

Step 1 Enable the global CAC function.

By default, the global CAC function is already enabled. You can run the **igmp bandwidthCAC** { **enable** | **disable** } command to change the setting.

Step 2 Configure the bandwidth of the multicast user.

Run the **igmp user add service-port** *index* **max-bandwidth** command to allocate the maximum bandwidth of the multicast user.

Step 3 Configure the bandwidth of the multicast program.

- Run the **igmp program add ip** *ip-addr* **bandwidth** command to configure the bandwidth of a single multicast program. The program bandwidth is an attribute of a multicast program, specifying the bandwidth requirement of the program being played.

- Run the **igmp bandwidth port** *frameid/slotid/portid* **max-bandwidth**{ *bandwidth* | **no-limit** } command to configure the program bandwidth of a physical port on a board. This command is available for only the GPON port. The default bandwidth of a port is 716800 kbit/s. Configuring the total program bandwidth for a single port is a way of traffic management, which helps avoid network congestion caused by the excessively-large total program bandwidth on a port. When the total program bandwidth of a port exceeds the value configured using the **igmp bandwidth port** *frameid/slotid/portid* **max-bandwidth**{ *bandwidth* | **no-limit** } command, subsequent programs ordered by users on this port cannot be played.

Step 4 Check whether the multicast bandwidth configuration is correct.

- Run the **display igmp config global** command to check the status of the global CAC function.
- Run the **display igmp program** command to query the bandwidth allocated to the multicast program.
- Run the **display igmp user** command to query the maximum bandwidth and the occupied bandwidth of the multicast user.

----End

Example

To enable bandwidth management for multicast users, set the user bandwidth to 10 Mbit/s when adding multicast user 0/2/1, and configure the program bandwidth to 1 Mbit/s when adding multicast program 224.1.1.1.

```
huawei(config)#btv
huawei(config-btv)#igmp bandwidthcAC enable
huawei(config-btv)#igmp user add port 0/2/1 max-bandwidth 10240
huawei(config-btv)#quit
huawei(config)#multicast-vlan 101
huawei(config-mvlan101)#igmp program add ip 224.1.1.1 bandwidth 1024
```

(Optional) Configuring Multicast Preview

Multicast preview is an advertizing method provided by carriers for ISPs. The purpose is to allow users to have an overview of a program so that the user can determine whether to request for the program. To protect the legitimate interests of ISPs, the duration, interval, and count of user previews need to be controlled.

Prerequisites

The program matching mode of the multicast VLAN must be the static configuration mode.

Context

The difference between program preview and normal program watching is that, after the user goes online, the duration of the preview is restricted. When the duration expires, the user goes offline. The user can request the program again only after the preview interval expires. The count by which the user can request the program within a day (the start time can be configured) is restricted by the preview count of the user.

Multicast preview parameters are managed through the preview profile. One program can be bound to only one preview profile, but one preview profile can be referenced by multiple programs.

NOTE
IPv6 multicast does not support the multicast preview function.

Table 18-24 lists the default settings of the multicast preview parameters.

Table 18-24 Default settings of the multicast preview parameters

Parameter	Default Value
Global multicast preview function	enable
Preview profile	Preview profile with index 0
Preview profile parameters	Maximum preview duration: 120s Maximum preview count: 8 Minimum interval between two previews: 120s
Time for resetting the preview record	4:00:00 am
Valid duration of multicast preview	30s

Procedure

In the global config mode, run the **btv** command to enter the BTV mode.

Step 1 Enable the global multicast preview function.

By default, the global multicast preview function is enabled. You can run the **igmp preview**{ **enable** | **disable** } command to change the setting.

Step 2 Configure the preview profile.

Run the **igmp preview-profile add** command to configure the preview profile, and set the parameters: maximum preview duration, maximum preview count, and minimum interval between two previews. The system has a default preview profile with index 0.

Step 3 Bind the program to the preview profile.

In the multicast VLAN mode, run the **igmp program add ip** *ip-addr* **preview-profile** *index* command to bind the program to be previewed to the preview profile so that the program has the preview attributes as defined in the preview profile. By default, the program is bound to the preview profile with index 0.

Step 4 Change the time for resetting the preview record.

Run the **igmp preview auto-reset-time** command to change the time for resetting the preview record. The preview record of the user remains valid within one day. On the second day, the preview record is reset. By default, the system resets the preview record at 4:00:00 a.m.

Step 5 Modify the valid duration of multicast preview.

Run the **igmp proxy recognition-time** or **igmp preview recognition-time** command to modify the valid duration of multicast preview. If the actual preview duration of the user is shorter than the valid duration, the preview is not regarded as a valid one and is not added to the preview count. By default, the valid duration of multicast preview is 30s.

NOTE

If you use **igmp proxy recognition-time** and **igmp preview recognition-time** commands to set the valid duration of multicast preview concurrently, the one set by the **igmp preview recognition-time** command takes effect.

Step 6 Run the **display igmp config global** command to check whether the values of the multicast preview parameters are correct.

----End

Example

To enable preview of multicast programs by using the system default preview profile, do as follows:

```
huawei(config)#btv
huawei(config-btv)#igmp preview enable
```

To enable preview of multicast programs, create preview profile 1, set the maximum preview time to 150s, the maximum preview count to 10, and apply this preview profile when adding program 224.1.1.1, do as follows:

```
huawei(config)#btv
huawei(config-btv)#igmp preview enable
huawei(config-btv)#igmp preview-profile add index 1 duration 150 times 10
huawei(config-btv)#quit
huawei(config)#multicast-vlan 101
huawei(config-mvlan101)#igmp program add ip 224.1.1.1 preview-profile 1
```

(Optional) Configuring Program Prejoin

A delay occurs when a user switches programs and the switching is processed by devices on the network. In program prejoin, the device receives the multicast stream of a program from the upper-layer multicast router to the uplink port before a user sends a request to join a program. In this manner, the multicast stream can be directly transmitted from the uplink port to the user port after the multicast user requests for a program, shortening the waiting time of the user for requesting for the program and shortening the delay in program switching. Therefore, user experience is enhanced.

Prerequisites

The program matching mode of the multicast VLAN must be the static configuration mode.

Context

Multicast program prejoin is the same as program request. The device plays the role of a user and sends the report packet for receiving in advance the multicast stream from the upper-layer multicast router to the upstream port.

After the prejoin function is enabled, if the upper-layer multicast router does not support static multicast entry forwarding, the unsolicited report function needs to be enabled so that the user

can request the program quickly. Generally, the upper-layer multicast router processes the user request by responding to the group-specific query and the general query.

NOTE

The configuration steps for IPv4 multicast and IPv6 multicast are similar but detailed commands are different. This topic describes the configuration steps for IPv4 multicast, and provides the configuration example for IPv6 multicast.

Table 18-25 lists the default settings of the prejoin parameters.

Table 18-25 Default settings of the prejoin parameters

Parameter	Default Value
Prejoin function	disable
Unsolicited report of IGMP packets	disable

Procedure

Enable the prejoin function.

Run the **igmp program add ip** *ip-addr* **prejoin enable** command to enable the prejoin function of a program. By default, the prejoin function is disabled.

Step 1 After the prejoin function is enabled, if the upper-layer multicast router does not support static multicast entry forwarding, the unsolicited report function needs to be enabled for IGMP packets.

- Run the **igmp program add ip** *ip-addr* **unsolicited enable** command to enable the unsolicited report function for IGMP packets. By default, the unsolicited report function is disabled.
- Run the **igmp unsolicited-report interval** command to modify the interval for unsolicitedly reporting IGMP packets. By default, the interval is 10s.

Step 2 Check whether the prejoin function is configured correctly.

- Run the **display igmp program** command to query the status of the prejoin function and the unsolicited report function.
- Run the **display igmp config vlan** command to query the interval for unsolicitedly reporting IGMP packets.

----End

Example

(IPv4)To enable the prejoin function when adding program 224.1.1.1 on multicast VLAN 101 for reduce the program waiting time of the users, do as follows:

```
huawei(config)#multicast-vlan 101
huawei(config-mvlan101)#igmp program add ip 224.1.1.1 prejoin enable
```

(IPv6)To enable the prejoin function when adding program ffff::1 on multicast VLAN 101 for reduce the program waiting time of the users, do as follows:

```
huawei(config)#multicast-vlan 101
huawei(config-mvlan101)#igmp program add ipv6 ffff::1 prejoin enable
```

(Optional) Configuring the Multicast Logging Function

Multicast log records the information about multicast programs watched by the multicast user, which provides a criterion for carriers to evaluate the viewership of multicast programs.

Prerequisites

If the syslog is used for reporting multicast logs, the syslog server must be properly configured.

If the syslog server is not configured, you can run the **igmp syslog disable** command to disable the multicast syslog reporting function to save system resources.

Context

Multicast logs have three control levels: multicast VLAN level, multicast user level, and multicast program level. The system generates logs only when the logging functions at the three levels are enabled.

When the user stays online for longer than the valid time for generating logs, the system generates logs in any of the following conditions:

- The user goes offline naturally, by force, or abnormally.
- The user is blocked or deleted.
- The program is deleted.
- The program priority is changed.
- The upstream port to which the program is bound changes.
- The VLAN of the upstream port to which the program is bound changes.
- The preview time is longer than the recognition time of the multicast preview, and the preview is valid.

NOTE

Recognition time of the multicast preview: When a user previews a program, the recognition time is used to determine whether the preview is valid. If the preview time of a multicast user is shorter than the recognition time, the preview is invalid and is not added to the preview count. In addition, no multicast log is generated. When a user watches a program, the recognition time is used to determine whether to generate a multicast log. If the watch time of a multicast user is shorter than the recognition time, no multicast log is generated.

- The user preview times out.
- The IGMP mode is switched.
- The bandwidth CAC is not passed.

When the user goes online, the system records only the online date and time. The system generates a complete log only when the user goes offline.

The MA5600T/MA5603T/MA5608T can report the multicast log to the log server in the syslog mode and the call detailed record (CDR) mode. By default, the MA5600T/MA5603T/MA5608T reports the log in the syslog mode.

- Syslog mode: Logs are reported to the syslog server in the form of a single log.
- CDR mode: Logs are reported to the log server in the form of a log file (.cvs). One log file contains multiple logs.

Table 18-26 lists the default settings of the multicast logging parameters.

Table 18-26 Default settings of the multicast logging parameters

Parameter	Default Value
Report mode of the multicast log	Syslog mode
Logging function at the multicast VLAN level	enable
Logging function at the multicast user level	enable
Logging function at the multicast program level	enable
Action report function of the multicast user	disable
Interval for automatically logging	2 hours
Minimum online duration for generating a valid log	30s
Parameters of the log report in the CDR mode	Report interval: 600s Maximum number of logs that can be reported each time: 200

Procedure

- Configure the parameters of the logging function of the multicast host.
  a. Enable the multicast logging functions.

     Multicast logs have three control levels: multicast VLAN level, multicast user level, and multicast program level. The system generates logs only when the logging functions at the three levels are enabled. By default, the three functions are enabled.

     - In the MVLAN mode, run the **igmp log** { **enable** | **disable** } command to configure the logging function at the multicast VLAN level.
     - In the BTV mode, run the **igmp user add service-port** *index* **log** { **enable** | **disable** } command to configure the logging function at the multicast user level.

       In the BTV mode, run the **igmp log record** { **user** | **mac** } command to configure the log record object. After the configuration, the device can record ordering action of users or multicast terminals identified by MAC addresses.
     - In the Multicast VLAN mode, run the **igmp program add ip** *ip-addr* **log** { **enable** | **disable** } command to configure the logging function at the multicast program level.

  b. Modify the interval for automatically logging.

     In the BTV mode, run the **igmp proxy log-interval** command to modify the interval for automatically logging. When the user stays online for a long time, the system generates logs at the preset interval. This is to prevent the problem that a log

is not generated when the user leaves the multicast group without sending a leave packet, which can affect the accounting. By default, the interval is two hours.

c. Modify the minimum online duration for generating a valid log.

In the BTV mode, run the **igmp proxy recognition-time** or **igmp log recognition-time** command to modify the minimum online duration for generating a valid log. If the user is in a multicast group (such as to preview a program) for shorter than the preset duration, the user operation is not regarded as a valid one and a log is not generated. A log is generated only when a user stays online for longer than the specified duration. By default, the minimum online duration is 30s.

NOTE

If you use **igmp proxy recognition-time** and **igmp log recognition-time** commands to set the minimum online duration for generating a valid log concurrently, the one set by the **igmp log recognition-time** command takes effect.

- (Optional) Configure the action report function of the multicast user.

  By default, the system uses the syslog mode to report multicast logs. You can run the **igmp user-action-report** command to configure the action report function of the multicast user. By default, the action report function of the multicast user is disabled.

  - **enable**: Enables the action report function of the multicast user. Logs are reported to the syslog server when a multicast user goes online and offline.
  - **disable**: Disables the action report function of the multicast user. Logs are reported to the syslog server only when a multicast user goes offline.

- Configure the function of CDR-mode log report.

  a. Configure the multicast log server and the data transmission mode for the CDR-mode log report.

  Run the **file-server auto-backup cdr** command to configure the active and standby multicast log servers.

  b. Enable the function of CDR-mode log report.

  In the BTV mode, run the **igmp cdr** { **enable** | **disable** } command to configure the function of CDR-mode log report. After the function is enabled, the MA5600T/MA5603T/MA5608T reports the local multicast logs to the multicast log server in the form of a file. After the function is disabled, the MA5600T/MA5603T/MA5608T reports each single log to the syslog server in the default syslog mode.

  c. Configure the parameters of the log report in the CDR mode.

  - In the BTV mode, run the **igmp cdr-interval** command to set the report interval. By default, the interval is 600s.
  - In the BTV mode, run the **igmp cdr-number** command to set the maximum number of logs that can be reported each time. When the number of the multicast logs in the CDR file reaches the preset value, the MA5600T/MA5603T/MA5608T reports the logs. By default, the maximum number is 200.

  d. Check whether the configuration is correct.

  - Run the **display file-server** command to query the configuration of the CDR multicast log server.
  - Run the **display igmp config global** command to query the status and other parameters of the function of CDR-mode log report.

----End

Example

To configure the multicast log to be reported to log server 10.10.10.1 in CDR mode through SFTP transmission, do as follows:

```
huawei(config)#file-server auto-backup cdr primary 10.10.10.1 sftp path temp user
  User Name(<=40 chars):aaa
  User Password(<=40 chars):      //Input the password.
huawei(config)#btv
huawei(config-btv)#igmp cdr enable
```

(Optional) Configuring the Maximum Number of Programs That Can Be Watched by the Multicast User

This topic describes how to configure the maximum number of programs that can be ordered by the multicast user at the same time. You can configure the maximum number of all programs that can be watched by the multicast user at the same time, or configure the maximum number of the different-level programs that can be watched by the multicast user.

Prerequisites

When you configure the maximum number of programs based on the program level, the program level must be configured at the same time and the programs must be configured statically.

Context

Table 18-27 lists the default settings of the max-program parameters.

Table 18-27 Default settings of the multicast max-program parameters

Parameter	Default Value
Maximum number of programs that can be watched by the multicast user	8
Grade of the multicast program	no-grade
Maximum number of programs of different priorities that can be watched by the multicast user	no-limit

Procedure

In the global config mode, run the **btv** command to enter the BTV mode.

Step 1 Configure the max-program of the multicast user.

Run the **igmp user add service-port** ***index*** **max-program** ***max-program-num*** command to set the maximum number of programs that can be watched by the multicast user.

Step 2 Configure the maximum number of programs of different priorities that can be watched by the multicast user.

Run the **igmp user watch-limit service-port *index*** command to set the maximum number of programs of different priorities that can be watched by the multicast user.

Step 3 Configure the grade of the multicast program.

In the multicast VLAN mode, run the **igmp program add ip *ip-addr* grade**command to configure the grade of a multicast program.

Step 4 Check whether the multicast max-program configuration is correct.

- Run the **display igmp user** command to query the maximum number of programs that can be watched and watching by the multicast user.
- Run the **display igmp program** command to query the grade of the multicast program.
- Run the **display igmp user extended-attributes service-port** command to query the maximum number of programs that can be watched and watching by the multicast user.

----End

Example

To set the user max-program to 10 when adding multicast service-port 0, set the user can watch 2 HDTV program, and configure the program grade to hdtv when adding multicast program 224.1.1.1, do as follows:

```
huawei(config)#btv
huawei(config-btv)#igmp user add service-port 0 max-program 10
huawei(config-btv)#igmp user watch-limit service-port 0 hdtv 2
huawei(config-btv)#quit
huawei(config)#multicast-vlan 101
huawei(config-multicast-vlan 101)#igmp program add ip 224.1.1.1 grade hdtv
```

(Optional) Configuring the Maximum Rate for Sending IGMP Packets

When the multicast and anti-DoS attack functions are enabled, the system reports DoS attack alarms and drops IGMP packets over the rate limit, if a user port sends such IGMP packets to the CPU. When the anti-DoS attack function is disabled, the system always sends IGMP packets to the CPU. If they are sent to the CPU at a rate higher than the limit, the system drops the IGMP packets that exceed the rate limit but does not report DoS attack alarms.

Prerequisites

The multicast function is enabled.

Context

When the multicast function is enabled, the system will always send the received IGMP packets to the CPU if no control is implemented over the process. Then if a user port receives a large number of IGMP packets, the IGMP packets for other users will not be processed and directly dropped instead.

Table 18-28 lists the default system settings.

Table 18-28 Default system settings

Parameter	Default Value
Anti-DoS attack function	disable
Maximum rate for sending IGMP packets to the CPU	63 pps

Procedure

Enable the anti-DoS attack function.

Run the **security anti-dos** { **enable** | **disable** } command to enable the anti-DoS attack function, which is disabled by default.

Step 1 Specify the maximum rate for sending IGMP packets to the CPU.

Run the **security anti-dos control-packet igmp rate** *frameid/slotid/portid* { **value** | **no-limit** } command to specify the maximum rate for sending IGMP packets to the CPU, which is 63 by default.

----End

Example

Example: Specify the maximum rate for sending IGMP packets to the CPU to 63 pps for user port 0/1/1, and enable the system to report the port to the blacklist if it sends IGMP packets over the rate limit to the CPU.

```
huawei(config)#security anti-dos enable
huawei(config)#security anti-dos control-packet igmp rate 0/1/1 20
```

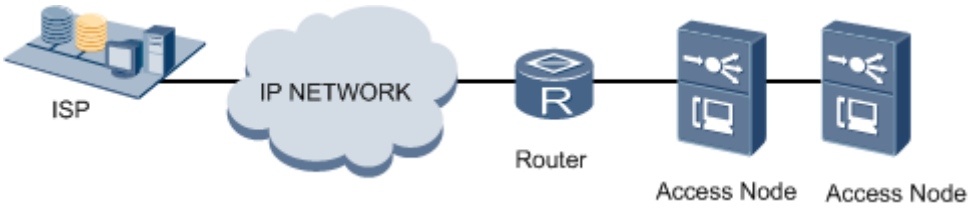
18.7.3 Configuring the Multicast Service in a Subtending Network

This topic describes how to configure the multicast service of the access device in a subtending network.

Application Context

Figure 18-55 shows the application context of the multicast service in a subtending network. When a subtended device needs to provision the multicast service, the subtending port on the subtending device needs to be configured as a multicast subtending port. In this way, the subtending device regards the subtended device as an IGMP user.

Figure 18-55 Application context of multicast service in a subtending network



Context

Default Configuration

Table 18-29 lists the default settings of the multicast service.

Table 18-29 Default settings of the multicast service

Feature	Default Settings
Multicast protocol	Disabled
IGMP version, includes multicast user version and multicast VLAN version.	V3
Multicast program configuration mode	Static configuration mode
Multicast bandwidth management	Enabled
Multicast preview	Enabled
Multicast log function	Enabled

Precaution

- The multicast program list of the subtending device must cover the multicast program list of the subtended device.
- In this network, the access device functions as a DSLAM, and the multicast VLANs of the subtending device and subtended device must be the same.

Procedure

The procedure for configuring the subtended device is the same as described in 18.7.2 Configuring the Multicast Service on a Single NE.

The procedure of configuring the subtending device is as follows:

1. For details on configuring the multicast service, see 18.7.2 Configuring the Multicast Service on a Single NE.
2. Configure the multicast subtending port.

   Run the **igmp cascade-port** *frameid/slotid/portid* command to configure the subtending port as the multicast subtending port. The multicast upstream port cannot be configured as a multicast subtending port.

3. Configure the mode for processing unknown multicast packets by the multicast subtending port.

   By default, the multicast subtending port transparently transmits the unknown multicast packets sent by lower-layer devices. This applies to the situation that the lower-layer devices may require the transparent transmission of unknown multicast packets. When multicast service is a priority, it is suggested to run the **igmp cascade-port** *frameid/slotid/portid* **mismatch** { **transparent** | **discard** } command to discard unknown multicast packets.

4. When the subtended device requires the quick leave function of the multicast user, run the **igmp cascade-port** *frameid/slotid/portid* **quickleave enable** command to enable the quick leave attribute on the multicast subtending port.

NOTICE

If the lower-layer device does not support the proxy of the IGMP leave packet, all the users requesting the program may go offline when a user requesting the same program goes offline. Therefore, when the quick leave attribute is enabled on the multicast subtending port, it is recommended that the lower-layer device use the IGMP proxy function, or switch to the IGMP snooping mode with the IGMP leave packet proxy function enabled.

18.7.4 Configuring the Multicast Service in an MSTP Network

This topic describes how to configure the multicast service in an MSTP network.

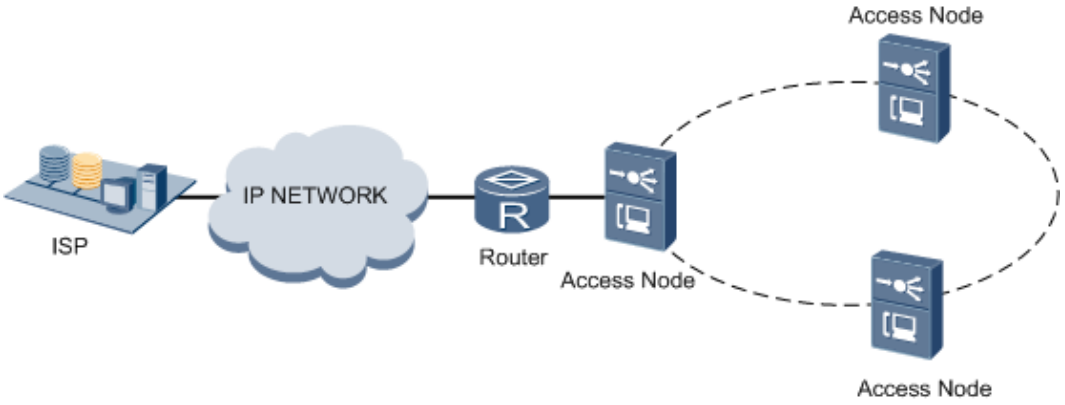
Prerequisites

Basic configurations for the MSTP network are complete. For details about configuring the MSTP network, see 19.6.3 Configuring the MSTP.

Application Context

Figure 18-56 shows the application context of the multicast service in an MSTP network. When the multicast service is provisioned in an MSTP ring network, the multicast upstream port and the subtending port need to be added to the multicast VLAN. According to the running result of the MSTP protocol, the multicast request packets are sent from the root port or the default port (when the device is a root bridge), and the other ports in the VLAN serve as subtending ports.

Figure 18-56 Application context of the multicast service in an MSTP network



Context

Default Configuration

Table 18-30 lists the default settings of the multicast service.

Table 18-30 Default settings of the multicast service

Feature	Default Settings
Multicast protocol	Disabled
IGMP version, includes multicast user version and multicast VLAN version.	V3
Multicast program configuration mode	Static configuration mode
Multicast bandwidth management	Enabled
Multicast preview	Enabled
Multicast log function	Enabled

Procedure

1. For details on configuring the MSTP ring network, see 19.6.3 Configuring the MSTP.
2. For details on configuring the multicast service, see 18.7.2 Configuring the Multicast Service on a Single NE.
3. Configure the MSTP multicast upstream port.

   When multicast service is provisioned in an MSTP network, the multicast upstream port needs to be set to the MSTP mode, and the default upstream port of the multicast VLAN can be specified. After the configuration is completed, multicast packets are forwarded by the root port or default port of the multicast VLAN.

   - Run the **igmp uplink-port-mode mstp** command to set the upstream port to the MSTP mode.

- Run the **igmp default uplink-port** command to specify the default upstream port of the multicast VLAN. When the upstream port is set to the MSTP mode and an MSTP root port is not available in the multicast VLAN, the multicast VLAN by default adopts the upstream port as the multicast upstream port.

4. Configure the multicast subtending port.

   Run the **igmp cascade-port** command to configure the subtending port as the multicast subtending port.

5. Configure multicast quick convergence in the case of an MSTP network topology change.

   Multicast quick convergence means that the device can quickly join the multicast group through a new upstream port when the MSTP network topology changes. The device can unsolicitedly send the new upstream port the IGMP join packet for an online program so that the device joins all the multicast groups. Or, the device can send the IGMP global leave packet to the upstream port. Then, the upper-layer querier sends a query packet for generating a new multicast forwarding tree.

   Run the **igmp send global-leave** command to enable the function of sending the IGMP global leave packet. When this function is enabled, the device sends the IGMP global leave packet to the upper-layer multicast router. When this function is disabled, the device sends the IGMP join packet to the upper-layer multicast router. By default, the function of sending the IGMP global leave packet is enabled.

18.7.5 Example of the xDSL Multicast Service

This topic describes how to configure the multicast video service.

Configuring the Multicast Video Service (Static)

Static program configuration for multicast services is fine-grained program configuration and management, with functions of program/user multicast bandwidth management, program preview, and program prejoin. This configuration mode applies to multicast services that have strict requirements on program and user management.

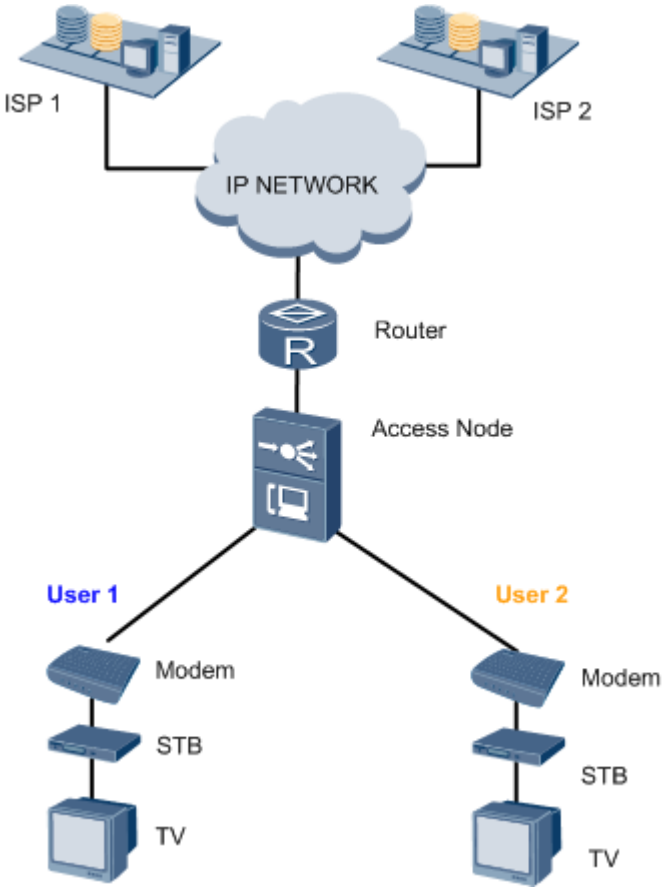
Service Requirements

- ISP 1 has two popular music programs, and ISP 2 has one popular music program and one popular video program. These programs can be ordered continuously and in time.
- The program ordering status can be recorded, collected, and reported for monitoring and charging.
- User 1 has purchased only the music program package from ISP 1, while user 2 has purchased only the video program package from ISP 2. Rights management is required for the two users.
- The package purchased by user 1 limits the maximum available multicast bandwidth to 10 Mbit/s, and the package purchased by user 2 limits the maximum available multicast bandwidth to 5 Mbit/s. Bandwidth restriction is required for the two users.

Figure 18-57 shows an example network of the multicast service.

User 1 and user 2 are connected to the access device in VDSL IPoE mode.

Figure 18-57 Example network of the multicast service



Data Plan

Table 18-31 lists the data plan for configuring the multicast video service with static program configuration.

Table 18-31 Data plan for configuring the multicast video service with static program configuration

Item	Data
Multicast VLAN	Multicast domain of ISP 1: VLAN 10 Multicast domain of ISP 2: VLAN 20 **NOTE** It is recommended that different multicast VLANs be planned for different ISPs.
Multicast upstream port	0/19/0
Multicast mode	IGMP proxy
Multicast protocol	IGMPv3
Multicast program	ISP 1

Item	Data
	• Program source address: 10.10.10.10 • program1: 224.1.1.1 • program2: 224.1.1.2 • Maximum program bandwidth: 3500 kbit/s • Program log reporting: enabled ISP 2 • Program source address: 10.10.10.11 • program3: 224.1.1.3 • program4: 224.1.1.4 • Maximum program bandwidth: 5000 kbit/s • Program log reporting: enabled
Right profile	Music program • Profile name: music • Program right: watch Video program • Profile name: movie • Program right: watch
VDSL port attribute	Working mode: PTM Bound line profile and alarm profile: default profiles (profile ID: 1) Port connected to user 1: 0/2/1 Port connected to user 2: 0/4/1
Multicast user	User 1: • Name of the bound right profile: music • Maximum multicast bandwidth: 10 Mbit/s • Multicast user log recording: enabled User 2: • Name of the bound right profile: movie • Maximum multicast bandwidth: 5 Mbit/s • Multicast user log recording: enabled

Procedure

Configure multicast VLANs and programs.

Configure smart VLAN 10 as the multicast domain of ISP 1, and smart VLAN 20 as the multicast domain of ISP 2.

1. Configure the protocol, multicast upstream port, and program list of multicast VLAN 10.

   Set the multicast VLAN to VLAN 10, multicast mode to IGMP proxy, multicast protocol to IGMPv3 (system default), multicast upstream port to port 0/19/0, statically configured

programs to 224.1.1.1 and 224.1.1.2, program source address to 10.10.10.10, and program bandwidth to 3500 kbit/s, and enable program log reporting.

```
huawei(config)#vlan 10 smart
huawei(config)#multicast-vlan 10
huawei(config-mvlan10)#igmp mode proxy
huawei(config-mvlan10)#igmp uplink-port 0/19/0
huawei(config-mvlan10)#igmp program add name program1 ip 224.1.1.1 sourceip
10.10.10.10
 bandwidth 3500 log enable
huawei(config-mvlan10)#igmp program add name program2 ip 224.1.1.2 sourceip
10.10.10.10
 bandwidth 3500 log enable
huawei(config-mvlan10)#quit
```

2. Configure the multicast protocol, multicast upstream port, and program list of multicast VLAN 20.

   Set the multicast VLAN to VLAN 20, multicast mode to IGMP proxy, multicast protocol to IGMPv3 (system default), multicast upstream port to port 0/19/0, statically configured programs to 224.1.1.3 and 224.1.1.4, program source address to 10.10.10.11, and program bandwidth to 5000 kbit/s, and enable program log reporting.

```
huawei(config)#vlan 20 smart
huawei(config)#multicast-vlan 20
huawei(config-mvlan20)#igmp mode proxy
huawei(config-mvlan20)#igmp uplink-port 0/19/0
huawei(config-mvlan20)#igmp program add name program3 ip 224.1.1.3 sourceip
10.10.10.11
 bandwidth 5000 log enable
huawei(config-mvlan20)#igmp program add name program4 ip 224.1.1.4 sourceip
10.10.10.11
 bandwidth 5000 log enable
```

Step 1 Configure right profiles named **music** and **movie** with the watch right, and bind the right profiles to the programs.

```
huawei(config-mvlan20)#btv
huawei(config-btv)#igmp profile add profile-name music
huawei(config-btv)#igmp profile profile-name music program-name program1 watch
huawei(config-btv)#igmp profile profile-name music program-name program2 watch
huawei(config-btv)#igmp profile profile-name music program-name program3 watch
huawei(config-btv)# igmp profile add profile-name movie
huawei(config-btv)#igmp profile profile-name movie program-name program4 watch
huawei(config-btv)#quit
```

Step 2 Activate the VDSL ports, and bind the ports to the line profile and alarm profile.

Bind VDSL port 0/2/1 and VDSL port 0/4/1 to the default line profile (line profile 1) and the default alarm profile (alarm profile 1).

```
huawei(config)#interface vdsl 0/2
huawei(config-if-vdsl-0/2)#deactivate 1
huawei(config-if-vdsl-0/2)#activate 1 template-index 1
huawei(config-if-vdsl-0/2)#alarm-config 1 1
huawei(config-if-vdsl-0/2)#quit
huawei(config)#interface vdsl 0/4
huawei(config-if-vdsl-0/4)#deactivate 1
```

```
huawei(config-if-vdsl-0/4)#activate 1 template-index 1
huawei(config-if-vdsl-0/4)#alarm-config 1 1
huawei(config-if-vdsl-0/4)#quit
```

Step 3 Configure multicast users.

1. Create the service channels of the multicast users.

   Create service port 100 on VDSL port 0/2/1, and service port 101 on VDSL port 0/4/1.

   ```
   huawei(config)#port vlan 10 0/19 0
   huawei(config)#port vlan 20 0/19 0
   huawei(config)#service-port 100 vlan 10 vdsl mode ptm 0/2/1 rx-cttr 2 tx-cttr 2
   huawei(config)#service-port 101 vlan 20 vdsl mode ptm 0/4/1 rx-cttr 2 tx-cttr 2
   ```

2. Configure the attributes of the multicast users.

   Configure multicast user 0/2/1 as the authentication type, with log reporting enabled, and with the maximum bandwidth 10 Mbit/s. Configure multicast user 0/4/1 as the authentication type, with log reporting enabled, and with the maximum bandwidth 5 Mbit/s.

   ```
   huawei(config)#btv
   huawei(config-btv)#igmp user add service-port 100 auth log enable max-bandwidth
   10240
   huawei(config-btv)#igmp user add service-port 101 auth log enable max-bandwidth 5120
   ```

3. Bind the multicast users to the right profiles.

   Bind VDSL user 0/2/1 to right profile **music**, and VDSL user 0/4/1 to right profile **movie**.

   ```
   huawei(config-btv)#igmp user bind-profile service-port 100 profile-name music
   huawei(config-btv)#igmp user bind-profile service-port 101 profile-name movie
   ```

4. Add the VDSL users to the multicast VLANs so that the VDSL users are multicast members.

   ```
   huawei(config-btv)#multicast-vlan 10
   huawei(config-mvlan10)#igmp multicast-vlan member service-port 100
   huawei(config-mvlan10)#multicast-vlan 20
   huawei(config-mvlan20)#igmp multicast-vlan member service-port 101
   huawei(config-mvlan20)#quit
   ```

Step 4 Save the configuration.

```
huawei(config)#save
```

----End

Result

- VDSL user 0/1/0 can watch the programs with IP addresses 224.1.1.1 and 224.1.1.2 that are provided by ISP 1, but VDSL user 0/1/0 cannot watch the program with IP address 224.1.1.3.
- VDSL user 0/2/0 can watch the program with IP address 224.1.1.4 that is provided by ISP 2.

Configuration File

```
vlan 10 smart
multicast-vlan 10
igmp mode proxy
igmp uplink-port 0/19/0
igmp program add name program1 ip 224.1.1.1 sourceip 10.10.10.10
 bandwidth 3500 log enable
igmp program add name program2 ip 224.1.1.2 sourceip 10.10.10.10
 bandwidth 3500 log enable
quit
vlan 20 smart
igmp mode proxy
igmp uplink-port 0/19/0
igmp program add name program3 ip 224.1.1.3 sourceip 10.10.10.11
 bandwidth 5000 log enable
igmp program add name program4 ip 224.1.1.4 sourceip 10.10.10.11
 bandwidth 5000 log enable
btv
igmp profile add profile-name music
igmp profile profile-name music program-name program1 watch
igmp profile profile-name music program-name program2 watch
igmp profile profile-name music program-name program3 watch
igmp profile add profile-name movie
igmp profile profile-name movie program-name program4 watch
quit
interface vdsl 0/2
deactivate 1
activate 1 template-index 1
alarm-config 1 1
quit
interface vdsl 0/4
deactivate 1
activate 1 template-index 1
alarm-config 1 1
quit
port vlan 10 0/19 0
port vlan 20 0/19 0
service-port 100 vlan 10 vdsl mode ptm 0/2/1 rx-cttr 2 tx-cttr 2
service-port 101 vlan 20 vdsl mode ptm 0/4/1 rx-cttr 2 tx-cttr 2
btv
igmp user add service-port 100 auth log enable max-bandwidth 10240
igmp user add service-port 101 auth log enable max-bandwidth 5120
igmp user bind-profile service-port 100 profile-name music
igmp user bind-profile service-port 101 profile-name movie
multicast-vlan 10
igmp multicast-vlan member service-port 100
multicast-vlan 20
igmp multicast-vlan member service-port 101
quit
save
```

Configuring the Multicast Video Service (Dynamic)

Dynamic program generation simplifies multicast service configuration and reduces maintenance cost, but does not support functions of program/user multicast bandwidth

management, program preview, and program prejoin. This configuration mode applies to multicast services that do not have strict requirements on program and user management or to those having programs and users managed on the upper-layer device.

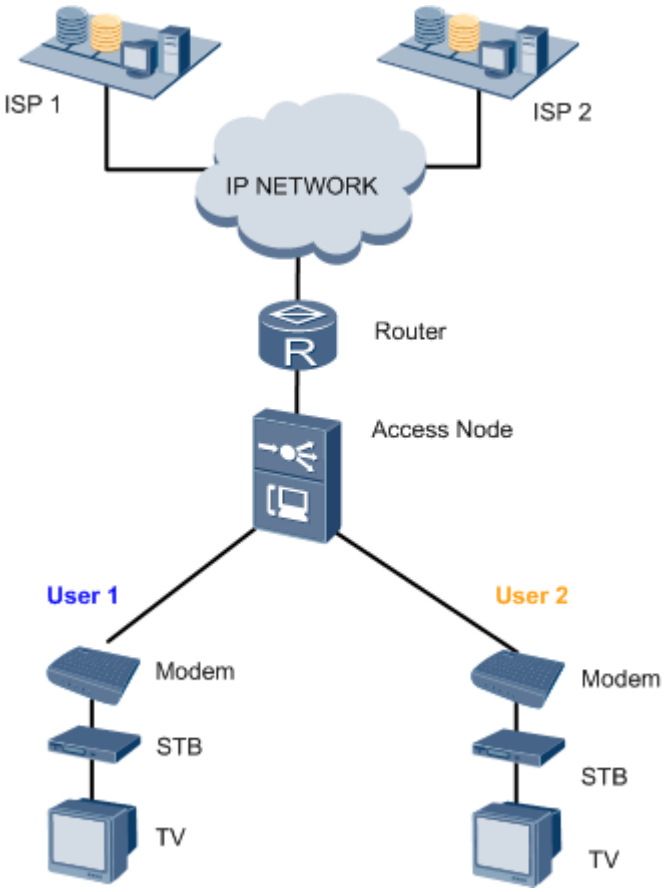
Service Requirements

- ISP 1 and ISP 2 each have an address segment corresponding to multicast programs, which do not require strict management and are dynamically updated according to the ordering situation.
- The service package purchased by user 1 allows user 1 to watch programs of ISP 1 only. The service package purchased by user 2 allows user 2 to watch programs of ISP 2 only.
- The ordering information about users can be recorded for monitoring and charging.

Figure 18-58 shows an example network of the multicast service.

User 1 and user 2 are connected to the access device in VDSL IPoE mode.

Figure 18-58 Example network of the multicast service



Data Plan

Table 18-32 lists the data plan for configuring the multicast video service with dynamic program generation.

Table 18-32 Data plan for configuring the multicast video service with dynamic program generation

Item	Data
Multicast VLAN	Multicast domain of ISP 1: VLAN 10 Multicast domain of ISP 2: VLAN 20 **NOTE** It is recommended that different multicast VLANs be planned for different ISPs.
Multicast upstream port	0/19/0
Multicast mode	IGMP proxy
Multicast protocol	IGMPv3
Multicast program	Dynamic generation Program address range of ISP 1: 224.1.1.1-224.1.1.2 Program address range of ISP 2: 224.1.1.3-224.1.1.4
VDSL port attribute	Working mode: PTM Bound line profile and alarm profile: default profiles (profile ID: 1) Port connected to user 1: 0/2/1 Port connected to user 2: 0/4/1
Multicast user	Multicast user log recording: enabled

Procedure

Configure multicast VLANs and programs.

Configure smart VLAN 10 as the multicast domain of ISP 1, and smart VLAN 20 as the multicast domain of ISP 2.

1. Configure the protocol, multicast upstream port, and program list of multicast VLAN 10.

   Configure multicast VLAN 10 with the dynamic program generation mode, and specify the range of the IP addresses of the programs that can be requested by the users in multicast VLAN 10 as 224.1.1.1 to 224.1.1.2. Multicast VLAN 10 adopts IGMP proxy, IGMP v3 (system default value), and multicast upstream port 0/19/0.

NOTICE

The multicast program configuration mode can be changed only when **igmp match mode** is set to **off**. However, setting **igmp match mode** to **off** will cause users to go offline. Therefore, it is recommended that the program configuration mode be planned beforehand.

```
huawei(config)#vlan 10 smart
huawei(config)#multicast-vlan 10
huawei(config-mvlan10)#igmp match mode disable
```

```
huawei(config-mvlan10)#igmp match group ip 224.1.1.1 to-ip 224.1.1.2
huawei(config-mvlan10)#igmp uplink-port 0/19/0
huawei(config-mvlan10)#igmp mode proxy
huawei(config-mvlan10)#quit
```

2. Configure the protocol, multicast upstream port, and program list of multicast VLAN 20.

   Configure multicast VLAN 20 with the dynamic program generation mode, and specify the range of the IP addresses of the programs that can be requested by the users in multicast VLAN 10 as 224.1.1.3 to 224.1.1.4. Multicast VLAN 20 adopts IGMP proxy, IGMP v3 (system default value), and multicast upstream port 0/19/0.

```
huawei(config)#vlan 20 smart
huawei(config)#multicast-vlan 20
huawei(config-mvlan20)#igmp match mode disable
huawei(config-mvlan20)#igmp match group ip 224.1.1.3 to-ip 224.1.1.4
huawei(config-mvlan20)#igmp uplink-port 0/19/0
huawei(config-mvlan20)#igmp mode proxy
huawei(config-mvlan20)#quit
```

Step 1 Activate the VDSL ports, and bind the ports to the line profile and alarm profile.

Bind VDSL port 0/2/1 and VDSL port 0/4/1 to the default line profile (line profile 1) and the default alarm profile (alarm profile 1).

```
huawei(config)#interface vdsl 0/2
huawei(config-if-vdsl-0/2)#deactivate 1
huawei(config-if-vdsl-0/2)#activate 1 template-index 1
huawei(config-if-vdsl-0/2)#alarm-config 1 1
huawei(config-if-vdsl-0/2)#quit
huawei(config)#interface vdsl 0/4
huawei(config-if-vdsl-0/4)#deactivate 1
huawei(config-if-vdsl-0/4)#activate 1 template-index 1
huawei(config-if-vdsl-0/4)#alarm-config 1 1
huawei(config-if-vdsl-0/4)#quit
```

Step 2 Configure multicast users.

1. Create the service channels of the multicast users.

   Create service port 100 on VDSL port 0/2/1, and service port 101 on VDSL port 0/4/1.

```
huawei(config)#port vlan 10 0/19 0
huawei(config)#port vlan 20 0/19 0
huawei(config)#service-port 100 vlan 10 vdsl mode ptm 0/2/1 rx-cttr 2 tx-cttr 2
huawei(config)#service-port 101 vlan 20 vdsl mode ptm 0/4/1 rx-cttr 2 tx-cttr 2
```

2. Configure the attributes of the multicast users.

   Enable the log reporting for multicast users 0/2/1 and 0/4/1. The authentication status and multicast bandwidth of the multicast users do not need to be configured.

```
huawei(config)#btv
huawei(config-btv)#igmp user add service-port 100 log enable
huawei(config-btv)#igmp user add service-port 101 log enable
huawei(config-btv)#quit
```

3. Add the VDSL users to the multicast VLANs so that the VDSL users are multicast members.

```
huawei(config)#multicast-vlan 10
huawei(config-mvlan10)#igmp multicast-vlan member service-port 100
huawei(config-mvlan10)#quit
huawei(config-btv)#multicast-vlan 20
huawei(config-mvlan20)#igmp multicast-vlan member service-port 101
huawei(config-mvlan20)#quit
```

Step 3 Save the configuration.

```
huawei(config)#save
```

----End

Result

- VDSL user 0/2/1 can watch the programs with IP addresses 224.1.1.1 and 224.1.1.2 that are provided by ISP 1.
- VDSL user 0/4/1 can watch the programs with IP addresses 224.1.1.3 and 224.1.1.4 that are provided by ISP 2.

Configuration File

```
vlan 10 smart
multicast-vlan 10
igmp match mode disable
igmp match group ip 224.1.1.1 to-ip 224.1.1.2
igmp uplink-port 0/19/0
igmp mode proxy
quit
vlan 20 smart
igmp match mode disable
igmp match group ip 224.1.1.3 to-ip 224.1.1.4
igmp uplink-port 0/19/0
igmp mode proxy
quit
interface vdsl 0/2
deactivate 1
activate 1 template-index 1
alarm-config 1 1
quit
interface vdsl 0/4
deactivate 1
activate 1 template-index 1
alarm-config 1 1
quit
port vlan 10 0/19 0
port vlan 20 0/19 0
service-port 100 vlan 10 vdsl mode ptm 0/2/1 rx-cttr 2 tx-cttr 2
service-port 101 vlan 20 vdsl mode ptm 0/4/1 rx-cttr 2 tx-cttr 2
btv
igmp user add service-port 100 auth log enable
igmp user add service-port 101 auth log enable
multicast-vlan 10
igmp multicast-vlan member service-port 100
multicast-vlan 20
```

```
igmp multicast-vlan member service-port 101
quit
save
```

18.8 Multicast Maintenance and Diagnosis

18.8.1 Multicast Emulation

A multicast emulation test remotely emulates an end user going online, and engineers query the real-time traffic of the multicast program to determine whether the multicast function is running properly.

Introduction

In multicast emulation, an access device remotely emulates an end user going online. Engineers query the real-time traffic of the multicast program to determine whether the multicast function is running properly.

Multicast emulation is used in acceptance tests or fault location. The following table lists the comparison between the multicast emulation test and traditional tests.

Table 18-33 Comparison between the multicast emulation test and traditional tests

Scenario	Task	Traditional Test	Multicast Emulation
Acceptance test	After an access device is installed and configured with data, a test engineer needs to check whether the multicast service has been provisioned to the access device successfully.	The test engineer visits the site where the access device is installed and uses an external tester or a portable computer to perform a multicast test on each port of the device.	The test engineer remotely logs in to the access device to perform a multicast emulation test and determines the service status based on the test results. **NOTE** A multicast emulation test cannot check the status of the line between an end user and an access device.
Fault location	The multicast service is abnormal, and a maintenance engineer needs to quickly locate the network segment of the fault to facilitate subsequent troubleshooting.	The maintenance engineer visits all sites where the access devices are installed and performs tests.	The maintenance engineer remotely logs in to an access device and performs a multicast emulation test to preliminarily determine the network segment of the fault. Based on the test results, the engineer diagnoses the fault cause and rectifies the fault.

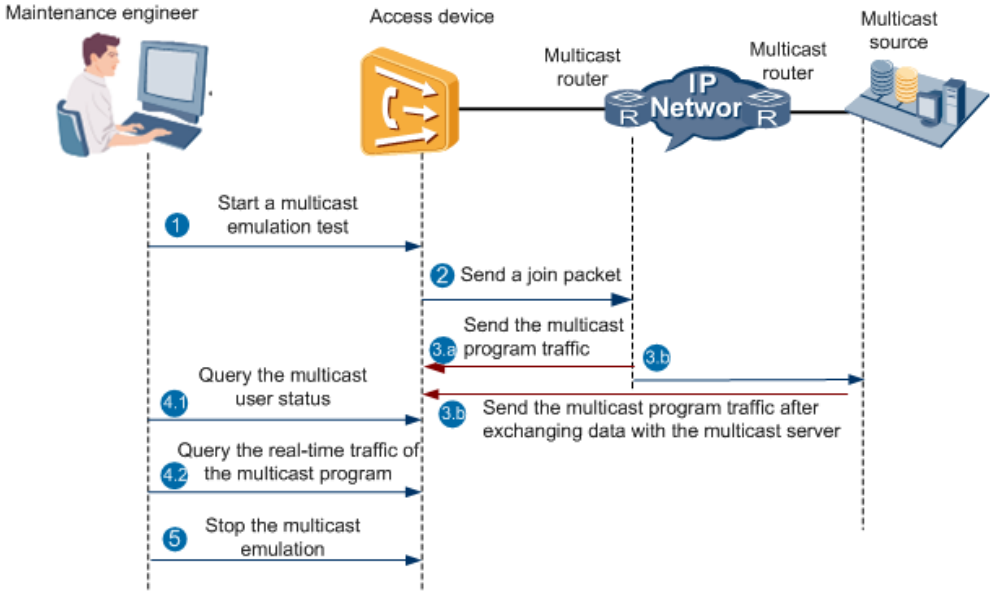
Reference Standards and Protocols

- RFC-2236: Fenner, W., "Internet Group Management Protocol, Version 2", RFC 2236, November 1997
- RFC 3376: B. Cain., "Internet Group Management Protocol, Version 3", RFC 3376, October 2002
- RFC 4607: H. Holbrook, "Source-Specific Multicast for IP", RFC 4607, August 2006

Principles

Figure 18-59 shows the principles of multicast emulation.

Figure 18-59 Principles of multicast emulation



1. A maintenance engineer remotely logs in to an access device and starts a multicast emulation test on a user port.

   The engineer sets parameters, such as the user information, program IP address, and multicast VLAN ID.
2. The access network device constructs a join packet and sends the packet to the multicast router for joining a multicast group.
3. The multicast router checks whether the multicast program traffic exists.
   - If the multicast program traffic exists, the multicast router sends the multicast program traffic to the access device.
   - If the multicast program traffic does not exist, the multicast source sends the multicast program traffic to the access device after exchanging data with the multicast router.
4. The maintenance engineer queries the status of the emulation user and the real-time traffic of the multicast program.
   - Checks whether the multicast emulation user is online to determine whether the user successfully orders the program.

- Checks the real-time traffic of the multicast program to determine whether the communication between the access device and the multicast source is normal.

5. After the multicast emulation is complete, the maintenance engineer stops the emulation manually using the CLI to release resources.

Usage Scenario

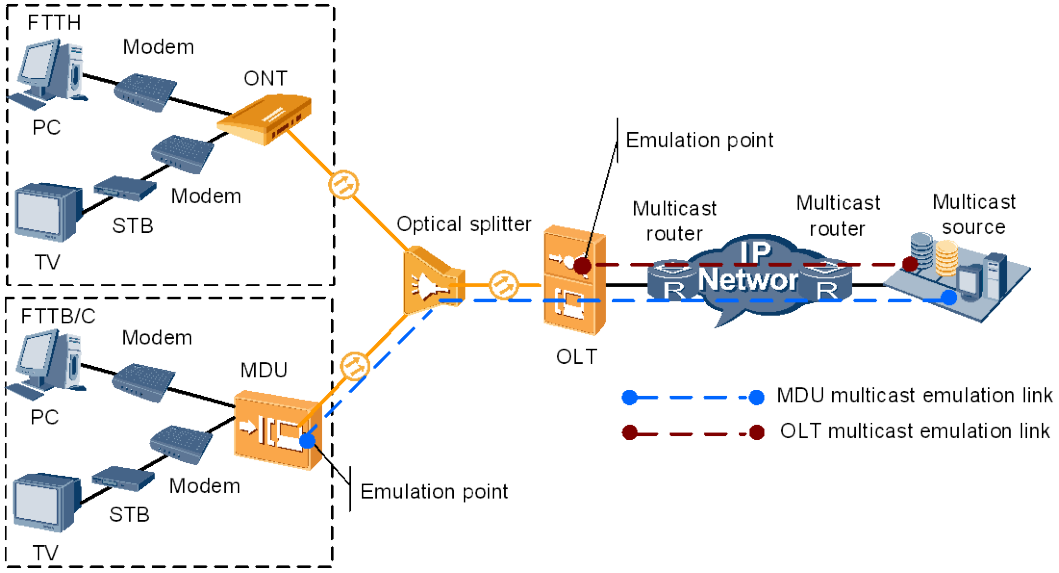
Context

On fiber to the x (FTTx) networks, access devices are located closer to user terminals and widely distributed. In a multicast emulation test, an access node emulates the multicast user to implement remote acceptance for services and locate faults, which reduces the O&M costs.

Scenario

Figure 18-60 shows the multicast emulation on a typical FTTx network.

Figure 18-60 Multicast emulation on a typical FTTx network



NOTE

In Figure 18-60, the MDU and ONT can provide the video on demand (VoD) service for a user through the PC. The MDU and ONT can also provide the BTV service using a set top box (STB).

In IPTV service acceptance or fault location:

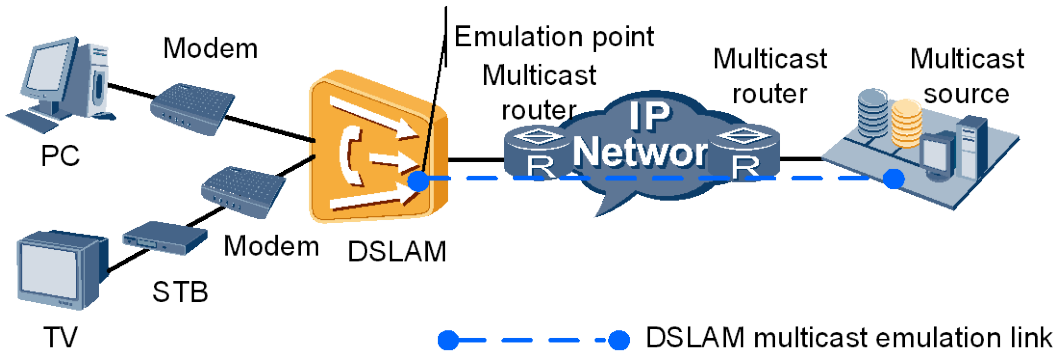
- For MDU multicast users, multicast emulation can be performed on the MDU in service acceptance.
- For ONT multicast users, multicast emulation can be performed on the OLT.

NOTE

ONTs do not support multicast emulation. To emulate an ONT multicast user, perform an emulation test on the PON board of the OLT.

Figure 18-61 illustrates the typical usage scenario of multicast emulation in DSLAM networking.

Figure 18-61 Typical usage scenario of multicast emulation in DSLAM networking



In IPTV service acceptance or fault location, multicast emulation can be performed on the DSLAM for the multicast users connected to the DSLAM.

Fault Location

After the multicast emulation, you can query the user status and multicast program's real-time traffic through the CLI. The following table lists troubleshooting suggestions based on the query results.

NOTE

Data configurations are seldom changed in daily operation and maintenance. Therefore, multicast faults are usually caused by hardware problems. Hardware must be checked prior to data configuration during fault location.

Table 18-34 Multicast emulation results and troubleshooting suggestions

Command	Results	Description	Troubleshooting Suggestions
display igmp user	The user status parameter **State** is **online**.	The multicast user can go online successfully.	• If the user can go online and the access device can communicate with the multicast source, the fault may be caused by a communication failure between the access device and the set-top box (STB). The reasons are as follows: – The hardware of the port on the access device is faulty. – The physical line between the access device and the modem is faulty. – The modem is faulty. – The STB is faulty. • If the user can go online

Command	Results	Description	Troubleshooting Suggestions
			but the traffic on the access device's uplink port is abnormal, the fault may be that the hardware connection between the access device and the upper-layer device (multicast router or multicast server) is incorrect, or the software configuration is incorrect. The common software configuration faults are as follows: – The remaining multicast bandwidth of the user is lower than the required bandwidth of the ordered program. – The number of programs watched by the multicast user reaches the upper limit so that the user cannot order a new program. – The multicast user does not have the permission to watch the program. – The program ordered is not in the MVLAN to which the multicast user belongs. – The multicast user does not have the permission to order certain types of programs (such as HDTV). – The number of programs at a level watched by the multicast user reaches the upper limit so that the user cannot order a new program at this level. – The rate configured in

Command	Results	Description	Troubleshooting Suggestions
			the traffic profile bound to the traffic stream is far lower than the bandwidth of the multicast program. – There are too many prejoined static programs, occupying too many bandwidths.

	The user status parameter **State** is **offline**.	The multicast user fails to go online and therefore fails to order a program in the emulation.	The case is more complicated than the scenario in which the user is **online**. Handling suggestions are as follows: Enable the multicast debugging function and start the multicast emulation again to check whether the access device receives the report packet from the user for ordering a program. Run the following commands to enable the multicast debugging function:
			`huawei(config)#terminal debugging`
			`huawei(config)#terminal monitor`
			`huawei(config)#debugging igmp service-port index`
			NOTE
			index is the ID of the multicast user's service port.
			• If the access device receives the report packet, the multicast link is normal but the access device fails to create a corresponding multicast entry. This is generally caused by incorrect multicast configurations on the access device.
			• If the access device does not receive the report

Command	Results	Description	Troubleshooting Suggestions
			packet, the multicast link fails. This is mainly caused by incorrect access device data configurations, faulty physical link between the access device and the modem, or hardware faults of terminals.
	The user status parameter **State** is **block**.	The multicast user is locked and therefore fails to order a program.	Run the **undo igmp user block** command to unblock the unblock the user.
display multicast flow-statistic	The multicast program's real-time traffic parameter **Multicast flow statistic result** is a small value or zero.	The program's traffic on the uplink port is small or zero.	The hardware connection between the access device and the upper-layer device (multicast router or multicast server) is incorrect, or the software configuration is incorrect. troubleshoot the fault based on the suggestions provided when the user can go online but the traffic on the access device's uplink port is abnormal.
	The multicast program's real-time traffic parameter **Multicast flow statistic result** is close to the program's bandwidth.	The device uplink port can communicate with the multicast source.	• If the user can go online and the access device can communicate with the multicast source, the fault may be caused by a communication failure between the access device and the STB. The reasons are as follows: – The hardware of the port on the access device is faulty. – The physical line between the access device and the modem is faulty. – The modem is faulty. – The STB is faulty. • If the access device can communicate with the multicast source but the user cannot go online,

Command	Results	Description	Troubleshooting Suggestions
			troubleshoot the fault based on the suggestions provided when the user status parameter **State** is **offline**.

Configuration

Context

- The configuration of the multicast port is correct.
- The multicast user for whom the multicast emulation test is performed has the rights to watch the configured multicast programs.

Procedure

Run the **igmp static-join** command to perform the multicast emulation test for the multicast user.

```
huawei(config)#btv
huawei(config-btv)#igmp static-join service-port 500
{ ip<K>|ipv6<K> }:ip
{ ip-addr<I><X.X.X.X> }:224.1.1.1
{ vlan<K> }:vlan
{ vlanid<U><1,4093> }:4002
{ <cr>|sourceip<K> }:
```

Step 1 Run the **display igmp user** command to query the status of the multicast user.

- If the multicast user is in **offline** state, the multicast user fails to request for programs.
- If the multicast user is in **online** state, the multicast user requests for programs successfully.
- If the multicast user is in **block** state, the multicast user is blocked. In this case, run the **undo igmp user block** command to unblock the user.

```
huawei(config)#display igmp user service-port 500
  User                   : 0/1/0
  State                  : online       //*The multicast user is online.
  Authentication          : auth
  Quick leave            : MAC-based
  IGMP flow ID           : 500
  Video flow ID          : 500
  Log switch             : enable
  Bind profiles          : 2
  IGMP version           : IGMP v3
  Current version         : IGMP v3
  IGMP IPv6 version       : IGMP IPv6 v2
  Current IGMP IPv6 version  : IGMP IPv6 v2
  Available programs      : 10
  Global leave           : disable
  User max bandwidth      : no-limit
```

```
 Used bandwidth(kbps)      : 0
 Used bandwidth
 to max bandwidth(%)       : -
 Total video bandwidth     : -
 Mcast video bandwidth     : -
 Bound profile list
 -----------------------------------------
 Index   Profile name        Program number
 -----------------------------------------
    0   profile0                    2
    1   Profile1                    8
 -----------------------------------------
 Total: 2
 Active program list
 ------------------------------------------------------------------------------
 Program name    VLAN  IP/MAC         State         Start time
 ------------------------------------------------------------------------------
 PROGRAM-0       4002  224.1.1.1      watching       2011-03-15
 ------------------------------------------------------------------------------
 Total: 10
```

Step 2 Run the **display multicast flow-statistic** command to query the real-time traffic of the programs that the multicast user requests for in the multicast emulation test.

- If the real-time traffic of the multicast programs is a smaller value or 0, the multicast source does not deliver multicast programs or the multicast service stream does not arrive at the device. That is, the communication between the device and the multicast source is abnormal.
- If the real-time traffic of the multicast programs approaches the bandwidth of the multicast programs, the multicast source delivers the multicast programs to the device. That is, the communication between the device and the multicast source is normal.

```
huawei(config)#btv
huawei(config-btv)#display multicast flow-statistic ip 224.1.1.1 vlan 4002
  Command is being executed, please wait...
  Multicast flow statistic result: 29600(kbps)   //The real-time traffic of multicast
program 224.1.1.1 is 29600 kbit/s. This indicates that the multicast source issues
multicast traffic.
```

Step 3 Run the **undo igmp static-join** command to stop multicast emulation.

```
huawei(config)#btv
huawei(config-btv)#undo igmp static-join service-port 500
{ ip<K>|ipv6<K> }:ip
{ ip-addr<I><X.X.X.X> }:224.1.1.1
{ vlan<K> }:vlan
{ vlanid<U><1,4093> }:4002
{ <cr>|sourceip<K> }:
```

----End

18.8.2 Video Quality Monitoring

Video quality monitoring enables an access device to remotely monitor affected video programs on configured monitoring points for fault demarcation.

Introduction

Background

After a fault occurs in the video service of a residential user, the user reports this fault to the carrier. Then, the carrier's maintenance personnel detect and rectify the fault. Before rectifying the fault, the maintenance personnel must demarcate the fault, involving the following cases:

- If multiple users report the fault related to an IPTV or VoD program within a period of time, and these users connect to different optical line terminals (OLTs), the IPTV server or core switching network may be faulty. In this case, the maintenance personnel only need to troubleshoot the IPTV server or core switching network.
- If only few users report the fault, and these users connect to the same OLT, the maintenance personnel must diagnose the video quality on the network for fault demarcation.

However, the maintenance personnel can obtain the video quality only by capturing video packets onsite. In the fiber to the home (FTTH) scenario, optical network terminals (ONTs) are located at user homes. Therefore, onsite packet capturing is not only time-consuming but also costly. In this case, an effective method of remotely diagnosing video quality is urgently required.

Overview

In video quality monitoring, monitoring points are configured on probes embedded into the OLT and ONTs. Each monitoring point is used for monitoring quality indicators of each video program. Maintenance personnel can remotely enable video quality monitoring and obtain monitoring results on the U2000 or OLT. Based on monitoring results obtained from each monitoring point, the maintenance personnel comprehensively identify fault points, thereby reducing fault locating costs.

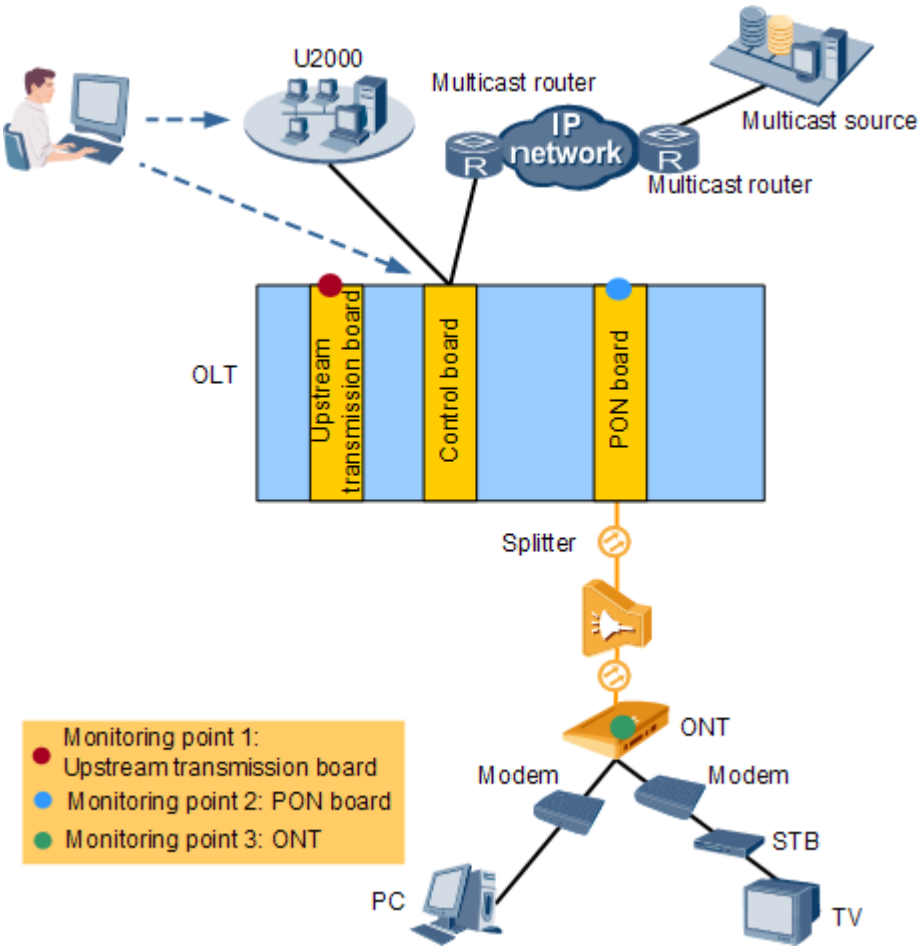
Compared with the video quality monitoring enabled using the OLT, the video quality monitoring enabled using the U2000 has the following advantages:

- The U2000 can store more video quality monitoring results than the OLT. In addition, the U2000 supports monitoring result dumping, meeting the requirement of long-term video quality monitoring. Major video quality-related faults occur randomly. Therefore, the identification of such faults requires long-term video quality monitoring data.
- The U2000 can display monitoring results in graphs. These trend graphs for each channel of program facilitate information obtaining.

Video quality monitoring points can be configured on Upstream transmission boards, PON boards of the OLT, and ONTs. In addition, the quality of both common videos and 4K videos can be monitored.

Figure 18-62 shows the application scenario of video quality monitoring.

Figure 18-62 Application scenario of video quality monitoring



Basic Concepts

Video quality monitoring involves the following quality indicators: media delivery index (MDI), mean opinion score for video (VMOS), and packet loss rate (PLR), which are described as follows:

MDI

MDI, defined in RFC 4445, is a quantitative number assigned to the quality of video streams transmitted over a network.

The MDI consists of two components: the delay factor (DF) and the media loss rate (MLR).

- DF is the maximum difference, observed at the end of each media data packet, between the arrival of media data and the forwarding of media data. This indicator reflects the delay and jitter of the tested video stream. A large jitter leads to a large DF value, which promotes high requirements on the buffer of the decoding device. Therefore, the smaller the DF value, the better. Recommended DF value range is 0-50 ms. Any DF value less than 200 ms is accepted.
- MLR specifies the number of media data packets lost per second. This indicator determines video quality. By measuring MLR, maintenance personnel can detect, identify, and trace media data packet loss on a network. According to WT126, the

maximum accepted MLRs for standard definition (SD)/VoD programs and high definition (HD) programs are 5 media data packets per 30 minutes and 5 media data packets per 240 minutes, respectively.

VMOS

VMOS is a subjective method of evaluating the quality of a video program.

A VMOS value is measured on a continuous scale of 1 to 5, representing the video quality of unsatisfactory, poor, fair, good, and excellent, respectively.

PLR

PLR specifies the number of IP packets lost per second. PLR is different from MLR, which specifies the number of media data packets lost in one sampling period. One IP packet contains seven media data packets.

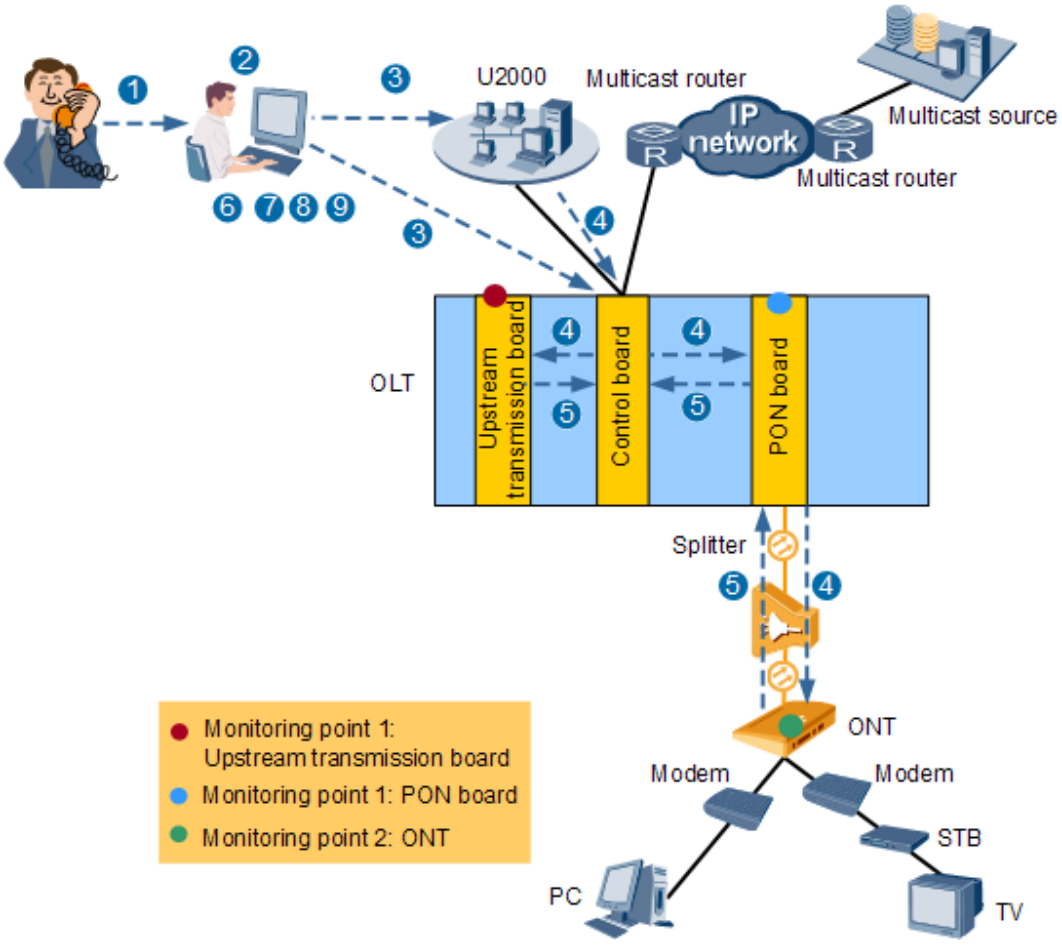
By measuring PLR, maintenance personnel can detect, identify, and trace IP packet loss on a network.

Principles

Implementation Process

Figure 18-63 shows the process of implementing video quality monitoring.

Figure 18-63 Process of implementing video quality monitoring



The process is as follows:

1. After a fault occurs in the video service of a residential user, the user reports this fault to the carrier.
2. The carrier's maintenance personnel exclude communication faults after checking the head end system or performing an Internet Group Management Protocol (IGMP) emulation test.
3. The maintenance personnel enable video quality monitoring using the OLT or U2000.

   Video quality monitoring can be enabled on Upstream transmission boards, PON boards or ONTs.

NOTE

If video quality monitoring is enabled on an uplink port, only the quality of inbound video streams can be monitored. If video quality monitoring is enabled on a non-uplink port, only the quality of outbound video streams can be monitored.

- To monitor the quality of an IPTV program, specify a multicast VLAN and multicast address.
- To monitor the quality of a VoD program, specify the source and destination IP addresses of this program, destination IP address, destination port, and transmission protocol.

4. The OLT or U2000 issues a command for enabling video quality monitoring to each monitoring point. After receiving this command, these monitoring points configure hardware based on the program information carried in the command. Then, the hardware captures video streams, adds timestamps to them, and sends the video streams to the CPU for monitoring indicator calculation.
5. The CPU sends calculated results to each monitoring point in real time, and the monitoring points periodically report these results to the control board of the OLT.
6. Video quality monitoring stops when the preset monitoring duration times out or the maintenance personnel disable this function.
7. The maintenance personnel obtain monitoring results from the OLT or U2000.
8. The maintenance personnel demarcate the fault based on monitoring results.
9. The maintenance personnel rectify the fault.

Fault Locating

After completing video quality monitoring, the maintenance personnel obtain monitoring results to learn program quality on each monitoring point for fault demarcation. Table 18-35 shows fault demarcation rules.

Table 18-35 Fault demarcation rules

Monitoring Point on a Upstream Transmission Board	Monitoring Point on a PON Board	Monitoring Point on an ONT	Conclusion
POK	POK	POK	The fault may occur on the upper-layer network or head end system. Detect the upper-layer metropolitan area network (MAN) of the OLT.
OK	POK	POK	The fault may occur between the PON board and the control board/upstream transmission board on OLT. Detect the inner OLT.
OK	OK	POK	The fault may occur between the PON board and the outbound port on the ONT (including the inner ONT).
OK	OK	OK	The fault may occur at user home, such as the Internet network or set top box (STB).

Configuring Video Quality Monitoring

Prerequisites

The carrier's maintenance personnel have excluded communication faults after checking the head end system or performing an IGMP emulation test.

Procedure

(Optional) Run the **video-quality-monitor** command to configure video quality monitoring indicator thresholds and a monitoring period.

Perform this step if default video quality monitoring indicator thresholds and monitoring period fail to meet network requirements. The monitoring indicators include DF, MLR, VMOS, and PLR.

Step 1 Run the **video-quality-monitor start** command to start a video quality monitoring instance.

You can configure one monitoring point on either a PON board or an ONT, or on both.

Step 2 (Optional) Run the **video-quality-monitor stop** command to stop the video quality monitoring instance.

Perform this step if you want to stop video quality monitoring before the monitoring duration times out.

Step 3 Run the **display video-quality-monitor result** command to query video quality monitoring results of the instance.

Demarcate the fault based on monitoring results.

1. Query the 5-minute VMOS availability ratio (AR) on each monitoring point.

NOTE

The 5-minute VMOS AR is the ratio of the number of available VMOS seconds to the total duration of 300s (5 minute x 60 seconds). When the VMOS value is greater than or equal to the fair threshold configured in the VMOS threshold profile, the VMOS value is available.

2. If the VMOS AR of a monitoring point is low, further query the detailed results of the video quality monitoring instance for fault identification.

----End

Example

NOTE

Terminal display significantly varies based on video quality monitoring instances. Therefore, the following example does not provide terminal display for queried commands. For details about the terminal display and parameter description, see the OLT Command Reference.

The U2000 provides monitoring results in graphs for simple information obtaining. Therefore, if the U2000 is available, use it to enable video quality monitoring and query monitoring results.

The following is an example configuration of the video quality monitoring feature:

- Excellent, good, and fair VMOS thresholds are 44, 33, and 22, respectively.
- Video quality monitoring points are configured on GPON board 0/2 and ONT 0 connected to port 0/2/0.
- The multicast VLAN is 100.
- The IP address of the multicast program is 224.1.1.1.

```
huawei(config)#video-quality-monitor vmos excellent-threshold 44 good-threshold 33
medium-threshold 22
huawei(config)#video-quality-monitor start service-board 0/2 ont 0/2/0 0
destination-ip 224.1.1.1 vlan 100
huawei(config)#display video-quality-monitor result board 0/2 ratio vmos
huawei(config)#display video-quality-monitor result ont 0/2/0 0 ratio vmos
huawei(config)#display video-quality-monitor result board 0/2 detail
huawei(config)#display video-quality-monitor result ont 0/2/0 0 detail
```

18.8.3 Common Multicast Maintenance Methods

This topic describes only the principles of fault diagnosis for multicast services. For details about troubleshooting (from fault symptom to troubleshooting procedure), see the *Troubleshooting* manual.

User Log

The device log records the program order history of users. The log includes the port to which a user is connected, IP address of the program group, multicast VLAN (MVLAN), time when a user starts watching a program, time when a user stops watching a program, and log mode (for example, watch, preview, idle, or preview threshold crossing).

A log is generated in any of the following scenarios:

- Normal channel switching. In such switching, the interval between receiving a leave packet to receiving a join packet is longer than the defined time.
- Ordering failure. The common causes of an ordering failure are as follows:
  - The user does not have the right to watch the program.
  - The maximum number of programs the user can watch concurrently is exceeded.
  - The bandwidth CAC fails.
- Daily preview threshold crossing.
- Quiet leaving. In quiet leaving, the user does not respond to the general query of the device.
- Long-time program watching. "Long-time watching" means that the watch time reaches the maximum duration configured in the system.
- Operations that cause a user to go offline, for example, deleting or blocking a user.

Logs can be queried according to different query criteria, including by user, by program, by a specified period with regard to a user, and by a specified period with regard to a program. If users need to learn only the log quantity, the log statistics function is recommended. This frees users from reading multiple pages of numerous logs that are generated after the device has been running for a long time.

Users can also use the log clearing function to delete unwanted old logs. Logs of all users or of a specified user be deleted.

Table 18-36 Commonly used multicast log commands

Command	Function
display igmp log	Query multicast logs.
display multicast failure-log	Query the multicast online and offline failure logs.

Command	Function
igmp log reset	Manually delete multicast logs.

Multicast Ping (Only OLT supports)

Using the CLI, multicast ping is a function with which a general query packet or group-specific query packet is sent to a specified multicast user or multicast cascading port. The version of such a packet is determined by the current version supported by the packet destination. After a report packet from the specified multicast user or multicast cascading port is received, the group IP address and source IP address of the report packet are displayed on the CLI. If no report packet is received within a specific duration, a timeout message is displayed on the CLI.

NOTE
Only xDSL/OPFA boards support this function.

IGMP Packet Statistics

A multicast program is ordered using the Internet Group Management Protocol (IGMP). Therefore, correctly sending, receiving, and processing IGMP packets is the prerequisite for successful program ordering. To facilitate fault locating in IGMP packet transmission/reception, the device supports three levels of IGMP packet statistics: global level, MVLAN level, and traffic stream level.

On the network side, the number of received IGMP query packets and number of sent IGMPv2/v3 join/leave packets can be queried based on MVLAN. According to the packet count, whether the upper-layer router is faulty can be determined.

On the user side, the number of received IGMPv2/v3 join/leave packets and number of sent IGMP query packets can be queried based on traffic stream. According to the packet count, whether a device in the home network is faulty can be determined.

Run the **display igmp statistic** command to query the multicast packet statistics.

Multicast Traffic Statistics Query

By querying multicast traffic statistics, users can determine whether multicast data reaches the ingress/egress of the device at the forwarding layer or whether packet loss occurs due to a low rate.

The device supports four query modes on the network side:

- A1: querying the number of sent/received multicast packets on an Ethernet port. Run the **display port statistics** command to query the statistics. The command output contains **Number of transmitted multicast frames** and **Number of received multicast frames**.
- A2: querying the ingress traffic (unit: kbit/s) of a specified multicast program (a pre-configured or dynamic program) or of a specified multicast upstream port. Run the **display multicast flow-statistic** command to query the traffic.
- A3: querying the number of required multicast packets (filtered by ACL) in the inbound direction of an Ethernet port.
- A4: querying the number of sent/received multicast frames on a PON port of an MDU that uses GPON upstream transmission. Run the **display gpon-port statistic** command

to query the number. The command output contains **Received multicast frames** and **Sent multicast frames**.

On the user side, the device supports three query modes for the GPON board:

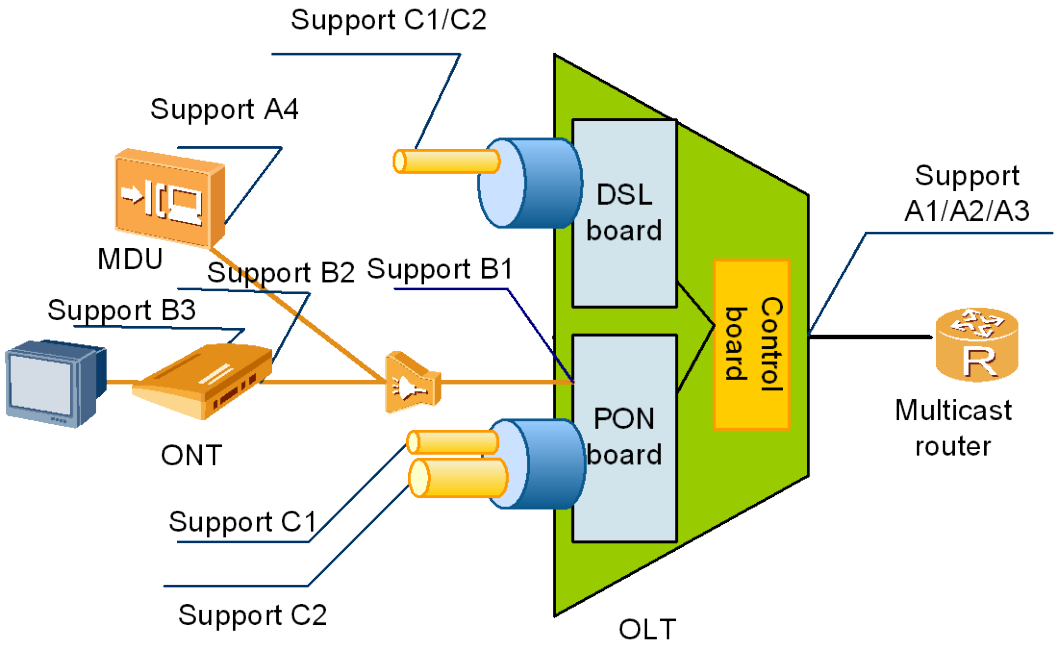
- B1: querying the number of sent/received multicast frames and number of sent bytes on a PON port. Run the **display statistics port ethernet** command to query the number. The command output contains **Received multicast frames** and **Sent multicast frames**.
- B2: querying the number of sent/received multicast frames of a specified ONT connected to an EPON board (The GPON board supports the command for this query mode but no multicast information is displayed.). Run the **display statistics ont** command to query the number. The command output contains **Received multicast frames** and **Sent multicast frames**.
- B3: querying the number of sent/received multicast frames on an Ethernet port of a specified ONT. Run the **display statistics ont-eth** command to query the number. The command output contains **Received multicast frames** and **Sent multicast frames**.

The following two methods can also be used to query the user-side multicast traffic statistics:

- C1: For the multicast traffic copied to service ports (for xDSL boards, P2P boards, and GPON boards that use multi-copy duplication), the traffic statistics (sent/received bytes) of a specified service port can be queried to obtain the multicast traffic statistics. This method is not suitable when a service port carries other services besides the multicast service.
- C2: For xDSL boards, P2P boards, or GPON/EPON boards that use single-copy duplication, the number of sent packets of queues on a specified port can be queried to obtain the multicast traffic statistics. This method is not suitable when queues on a port carry other services besides the multicast service.

Figure 18-64 shows the points where multicast traffic can be queried.

Figure 18-64 Points of multicast traffic query



Program Ordering Behavior Analysis

Compared with traditional TV services, in the case of IP multicast service, users' order behavior can be measured and analyzed at a finer grain, such as statistics measurement of hottest programs, analysis of user interest, and peak hours of program ordering. For such purposes, the device needs to precisely record the order behavior of each user in the form of logs and output the content of the logs through an open interface. According to different output modes, the device supports two log transfer modes: by CDR or by syslog (RFC 3164). The formats of the two modes are the same. For details, see "Charging Mode." The following table lists the pros and cons of the two modes.

Table 18-37 Pros and cons of the two log transfer modes

	Pro	Con
CDR	Reliable transfer. TFTP, FTP, or SFTP can be selected as the transfer protocol. **NOTE** SFTP is recommended.	Logs are reported to the file server only when specified reporting conditions are met (the reporting interval expires or the number of logs reaches the reporting threshold).
syslog	Timely report. Once a log is generated, it is uploaded to the syslog server.	Unreliable transfer. Syslog adopts the UDP protocol.

18.9 Reference Documents

The reference standards and protocols of this feature are as follows:

Table 18-38 Reference standards and protocols

Standard NO.	Standard Description	Application Scope
TR101	Technical Report DSL Forum, TR-101 Migration to Ethernet-Based DSL Aggregation, April 2006	IPv4 and IPv6 Multicast
TR156	Technical Report Broadband Forum, TR-156 Using GPON Access in the context of TR-101, December 2008	IPv4 and IPv6 Multicast
RFC 1112	Deering, S., "Host Extensions for IP Multicasting",STD 5, RFC 1112, August 1989	IPv4 and IPv6 Multicast
RFC-2236	Fenner, W., "Internet Group Management Protocol, Version 2", RFC 2236, November 1997	IPv4 and IPv6 Multicast
RFC 3376	B. Cain., "Internet Group Management Protocol, Version 3 ", RFC 3376,October 2002	IPv4 and IPv6 Multicast
RFC 3569	S. Bhattacharyya, "An Overview of Source-Specific Multicast (SSM)", RFC 3569, July 2003	IPv4 and IPv6 Multicast

Standard NO.	Standard Description	Application Scope
RFC 4601	B. Fenner, "Protocol Independent Multicast - Sparse Mode (PIM-SM): Protocol Specification (Revised)", RFC4601, August 2006	IPv4 and IPv6 Multicast
RFC 4604	H. Holbrook, "Using IGMPv3 and MLDv2 for Source-Specific Multicast", RFC 4604, August 2006	IPv4 and IPv6 Multicast
RFC 4605	B. Fenner, "IGMP/MLD Proxying", RFC 4605, August 2006	IPv4 and IPv6 Multicast
RFC 4607	H. Holbrook , "Source-Specific Multicast for IP", RFC 4607, August 2006	IPv4 and IPv6 Multicast
RFC 4541	M. Christensen, "Considerations for IGMP and MLD Snooping Switches", RFC 4541, May 2006	IPv4 and IPv6 Multicast
RFC 2710	S. Deering,"Multicast Listener Discovery (MLD) for IPv6",RFC 2710, October 1999	IPv6 Multicast
RFC 3810	R. Vida, Ed.,"Multicast Listener Discovery Version 2 (MLDv2) for IPv6", RFC 3810, June 2004	IPv6 Multicast

19 Network Protection Features

About This Chapter

This topic describes network protection features implemented by the system.

19.1 Network Protection Overview

The MA5600T/MA5603T/MA5608T provides a comprehensive network protection solution, including uplink protection, device protection, and PON downlink protection, enhancing reliability of the entire network.

To ensure reliability of the access network, the MA5600T/MA5603T/MA5608T supports the following end-to-end (E2E) network protection solution. BFD supports fast fault detection for both static and dynamic routes, ensuring the reliability of Layer 3 route forwarding, meeting carrier-class switching requirement (50 ms).

Figure 19-1 Protection scheme in an access network

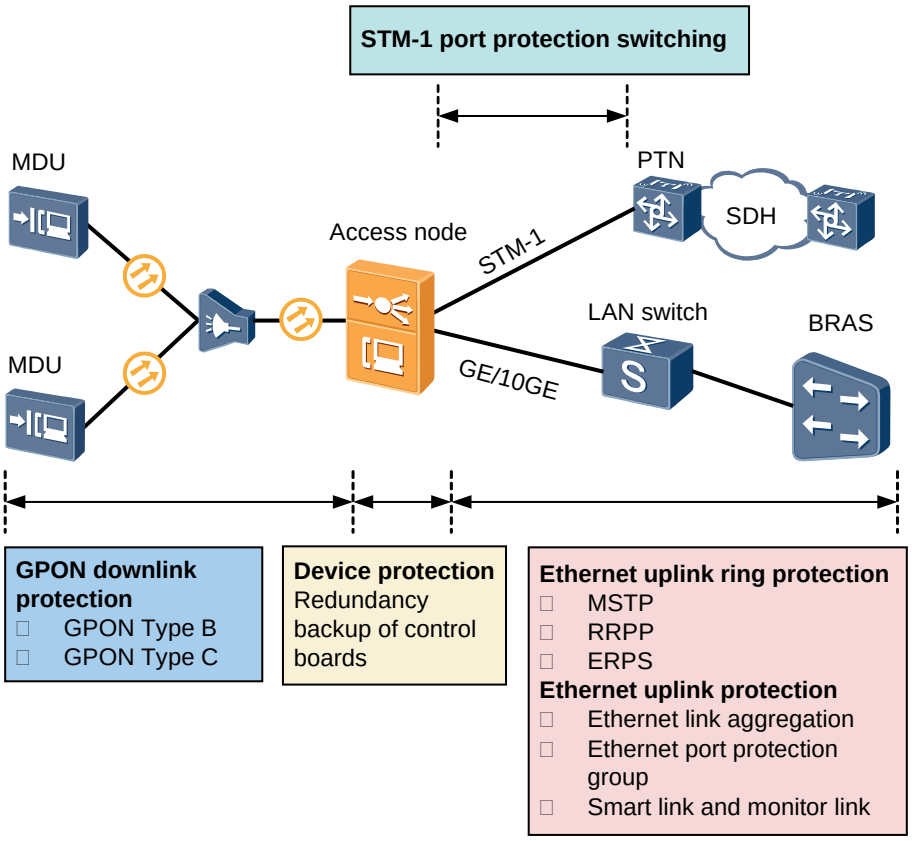


Table 19-1 Network protection scheme

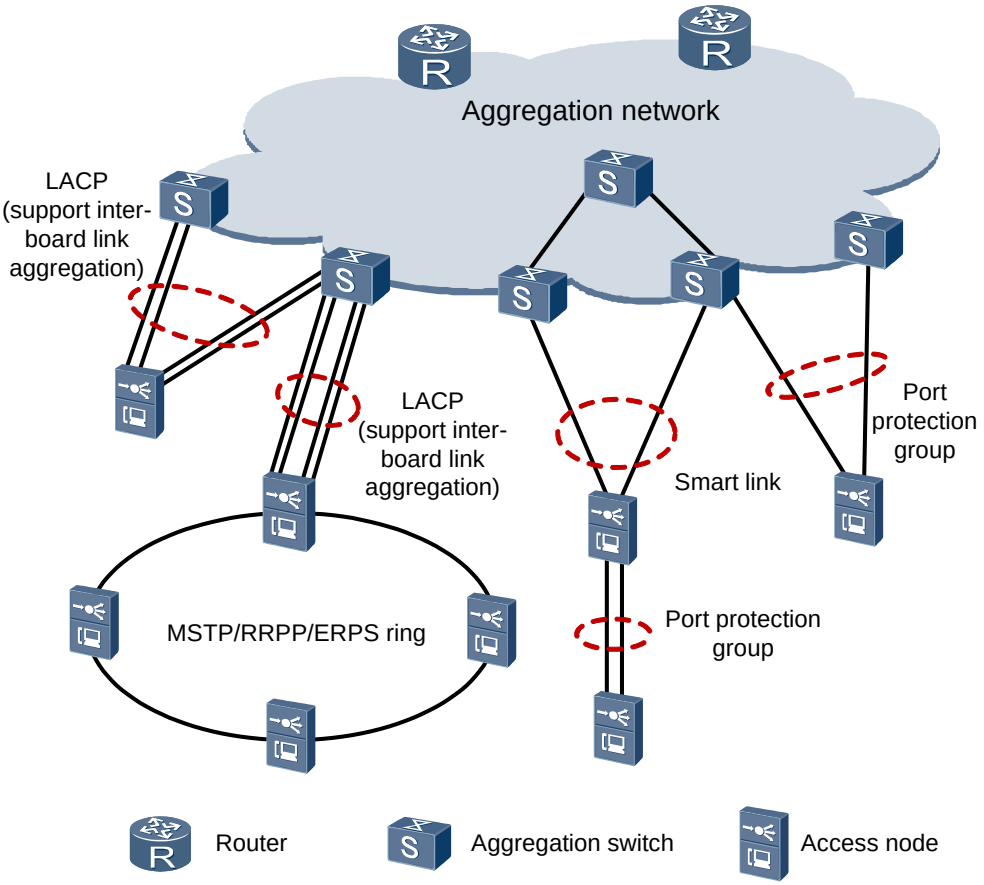
Protected Object	Protection Scheme	Huawei Proprietary Protocol or Not	Function and Feature
Device hardware	Redundancy backup of control boards • Active/standby mode • Load-sharing mode	No	An Ethernet port protection group can be configured only after active/standby redundancy backup of control boards is configured. • The active/standby mode improves reliability of the device and services. • The load-sharing mode doubles bandwidths and improves data forwarding performance of the device.
Ethernet uplinks or	Ethernet link aggregation	No	Increases uplink bandwidths, achieving load sharing. Members in the aggregation group

Protected Object	Protection Scheme	Huawei Proprietary Protocol or Not	Function and Feature
cascade links	• Manual aggregation • Link aggregation		back up for each other, enhancing uplink reliability. Manual aggregation is required if the MA5600T/MA5603T/MA5608T is connected to a device that does not support LACP.
	Ethernet port protection group	No	Supports uplink port protection, which can be used with Ethernet link aggregation.
	Smart link and monitor link	Yes	Smart links improve dual-uplink protection switching performance, which is faster than aggregation group switching or protection group switching. As a supplementary to the smart link solution, the monitor link solution is introduced to monitor uplinks. This improves the backup function of the smart link solution. This protection scheme is supported only when the MA5600T/MA5603T/MA5608T is connected to a Huawei device.
Ethernet uplink ring protection	MSTP	No	MSTP prunes a loop network to a loop-free tree network to avoid proliferation and infinite loop of packets in the loop network. MSTP is compatible with STP and RSTP. Furthermore, MSTP remedies drawbacks of STP and RSTP. The convergence time is subject to the network topology.
	RRPP	Yes	RRPP is a link-layer protocol dedicated to Ethernet ring protection. RRPP provides fast topology convergence within 50 ms and supports a convergence duration independent from the network topology. This protection scheme is supported only when the MA5600T/MA5603T/MA5608T is connected to a Huawei device.
	ERPS	No	ERPS is a ring network protocol defined in Recommendation ITU-T G.8032. Interoperation is achieved if all devices participating in the ring network support ERPS. In addition, ERPS supports fast

Protected Object	Protection Scheme	Huawei Proprietary Protocol or Not	Function and Feature
			convergence, meeting carrier-class reliability requirements.
STM-1 uplink protection	STM-1 port protection switching	No	STM-1 port protection improves reliability of uplinks when the STM-1 ports are connected to an SDH device to provide the TDM service.
GPON downlink protection	GPON type B	No	Type B provides redundancy for OLT's GPON ports and backbone fiber. Compared with type C, type B requires a lower cost but the implementation is more difficult.
	GPON type C	No	Type C provides redundancy for ONU's GPON ports, backbone fibers, optical splitters, and distribution optical fibers. Compared with type B, type C provides higher reliability.

Multiple Ethernet uplink network protection schemes can be used in combination, as shown in the following figure.

Figure 19-2 Ethernet uplink network protection



19.2 Redundancy Backup of Control Boards

An MA5600T/MA5603T/MA5608T configured with two control boards can work in the active/standby mode or load-sharing mode. The active/standby mode improves the reliability of the system and services. The load-sharing mode doubles the system bandwidth and enhances data forwarding performance.

19.2.1 Introduction to Control Board Redundancy Backup

Definition

In this feature, two control boards, one working as the active and the other working as the standby, are configured on an MA5600T/MA5603T/MA5608T to back up each other. When two control boards are powered on at the same time, the control board in the slot with a smaller slot ID functions as the active control board by default.

The redundancy backup of control boards benefits customers with the following capabilities:

- Quick switching between active and standby control boards
- Uninterrupted data forwarding during active/standby switchover
- Concurrent upgrade of the active and standby control boards

The two control boards can work in the active/standby mode or load-sharing mode. You can run the **system ex-mode** command to switch the working mode of the control boards.

NOTICE

Mode switching resets the system, clears configuration data (including saved configuration data), and interrupts the in-band network management channel. Therefore, you are advised to determine the working mode of control boards before site deployment and avoid switching the working mode afterwards.

Active/standby mode

As the core of the MA5600T/MA5603T/MA5608T, the active control board communicates with external devices and implements functions of internal modules of the system. The standby control board does not communicate with external devices and only serves as a backup of the active control board. During its operation, the active control board backs up all static configurations and some dynamic configurations to the standby control board to keep data synchronized between the two boards.

Redundancy backup of control boards protects services against a control board failure. If two control boards are configured, services can be switched to the standby control board when the active control board fails. Any of the following conditions triggers a switchover between the active and standby control boards:

- Active control board failure. In this case, the system performs an active/standby switchover automatically.
- System upgrade. In this case, the operator resets the control boards and performs the active/standby switchover manually.
- Board replacement or annual maintenance. In this case, the operator performs the active/standby switchover manually.

Load-sharing mode

When the two control boards work in load sharing mode, redundancy backup improves reliability of services as well as doubling bandwidth and enhancing data forwarding performance.

- On the forwarding plane, the active and standby control boards share loads. Both boards forward data.
- On the control plane, the two control boards work in the active/standby mode. The CPU on the active control board manages the system and controls data forwarding while the CPU on the standby control board is in the standby state.

19.2.2 Control Board Redundancy Backup Principle

Active/Standby Mode

Switchover Modes

- **Automatic active/standby switchover**

When the active control board fails, the system automatically performs an active/standby switchover. Specifically, the system resets the active control board, and the standby control board functions as the new active control board. During the whole process, the system continues processing data and services are not interrupted.

- **Manual active/standby switchover**

  When you need to replace the active control board or upgrade the system software, manually perform an active/standby switchover using one of the following methods:

  - Run the **system switch-over** command.
  - Run the **reboot** command to reset the active control board.
  - Press the RESET button on the active control board to reset the board.
  - Remove the active control board.

NOTE

During system running, if an active/standby switchover is performed by removing the active control board or pressing the RESET button, the hardware may be faulty or software data is not synchronized, leading to service failures. In this case, it is not recommended to remove the active control board or press the RESET button to perform an active/standby switchover. Specially, do not remove or install the active control board frequently in a short period of time.

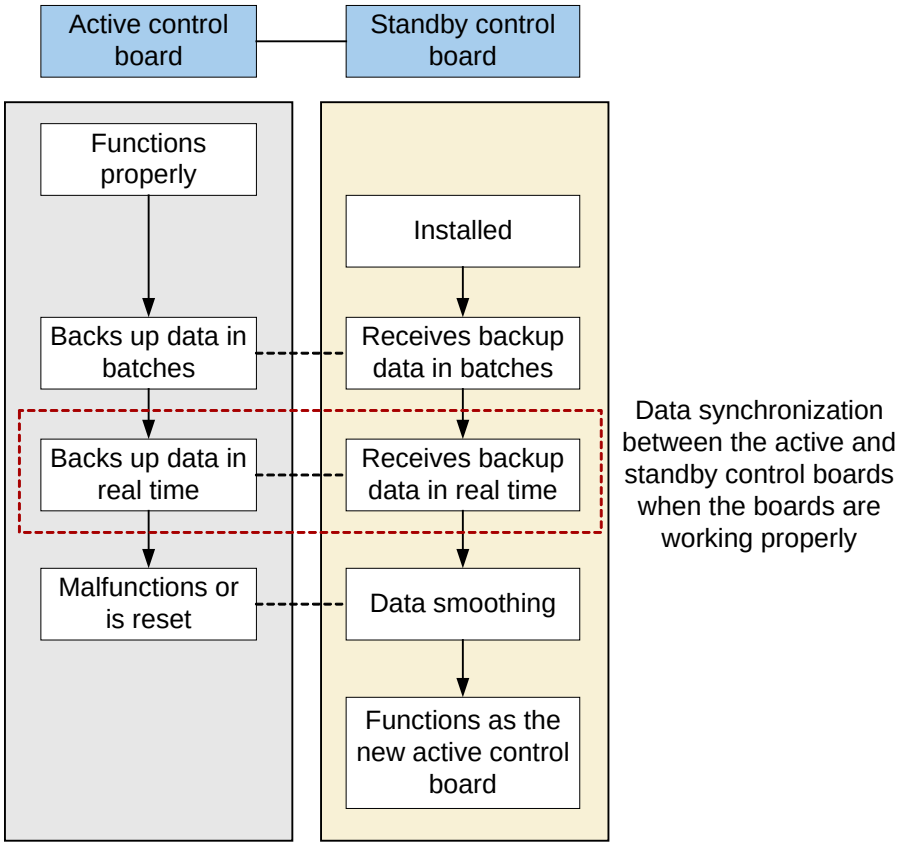
Communication Between the Active and Standby Control Boards

When two control boards are configured, the active control board is working and the standby control board is in the standby state. The standby control board exchanges information only with the active control board. The maintenance Ethernet port (designated with the ETH silkscreen) on the standby control board does not receive configuration commands. The serial port (designated with the CON silkscreen) on the standby control board receives configuration commands and forwards the commands to the active control board. Then the active control board issues the configurations to service boards.

The active and standby control boards achieve data consistency between them using high availability (HA) mechanisms such as batch backup and real-time backup. When learning that the active control board fails or receiving an active/standby switchover command, the standby control board takes over as the active control board and reports switchover events.

Figure 19-3 shows the state transition between the active and standby control boards in active/standby mode.

Figure 19-3 State transition between the active and standby control boards



- **Batch backup**

  After the standby control board is installed, the active control board automatically backs up data of all modules to the standby control board in batches.

- **Real-time backup**

  After batch backup is complete, the system enters the real-time backup state. In this state, the active control board backs up data to the standby control board in real time. When both control boards function properly, the system stays in the real-time backup state.

  Real-time backup maintains data consistency (also called data synchronization) between the active and standby control boards. The following data needs to be synchronized:

  - Configuration data: core data, also the static data, running in the system, including all configurations manually issued by users and data that is generated based on user configurations in system initialization and ensures normal system running.
  - Basic operating data: data that may change rarely during system running, including the device status (such as board status and port status), operation logs, and alarms. The basic operating data changes when the system status changes, for example, a board is faulty or a port connection status changes.
  - Dynamic service data: data that changes in real time during system running, including data generated by a call service (such as a PPPoE dialup) and data changing quickly (such as an ARP entry change). The dynamic service data takes a majority of data to be synchronized, and it is the most difficult to be completely synchronized. Data smoothing is mainly performed for such data.

You can run the **display data sync state** command to query the data synchronization status (complete synchronization or incomplete synchronization).

- When all data on the standby control board is completely synchronized with that on the active control board, it is called complete synchronization.

  An active/standby switchover performed when data is completely synchronized is called a **normal switchover**. A normal switchover does not interrupt services. A normal switchover can be performed only after data (including configuration data, basic operating data, and dynamic service data) is completely synchronized and the cyclic redundancy check (CRC) results on the active and standby control boards are the same.

- When data on the standby control board is not completely synchronized with that on the active control board, it is called incomplete synchronization.

  An active/standby switchover performed when the data is not completely synchronized is called a **forced switchover**. Table 19-2 lists three conditions of incomplete synchronization and whether forced switchover is supported in each condition.

Table 19-2 Active/standby switchover in the case of incomplete synchronization

Data Incompletely Synchronized	Supports Forced Switchover by Commands	Supports Forced Switchover by Board Reset (Manually or Using Commands) or Board Removal	Switchover Results
Configuration data	No	Yes	The system resets and services are interrupted. Forced switchover is not recommended.
Basic data	No	Yes	The system does not reset but the service board may reset and some services are interrupted. Forced switchover is not recommended.
Some dynamic service data	Yes	Yes	Data that is not synchronized to the standby control board is lost after the switchover, but services are not affected, and data about connections, alarms, and logs is not lost.

- **Data smoothing**

  In the real-time backup state, if an active/standby switchover occurs, the standby control board will be promoted to the active role. Before the standby control board becomes active, modules on the standby control board collect and synchronize data from service boards. The data collection and synchronization process is called data smoothing. During data smoothing, modules on the standby control board actively communicate with service boards to confirm and synchronize hardware status, link layer status, and configuration data. In this manner, data and status information are consistent through the entire system so that the system can run normally after the switchover.

The data smoothing duration is very short. After data smoothing is complete, the standby control board works as the new active control board.

Communication Between Service Boards and Control Boards

Service boards are connected to the active and standby control boards through buses. Data of service boards and data of the standby control board come from the active control board.

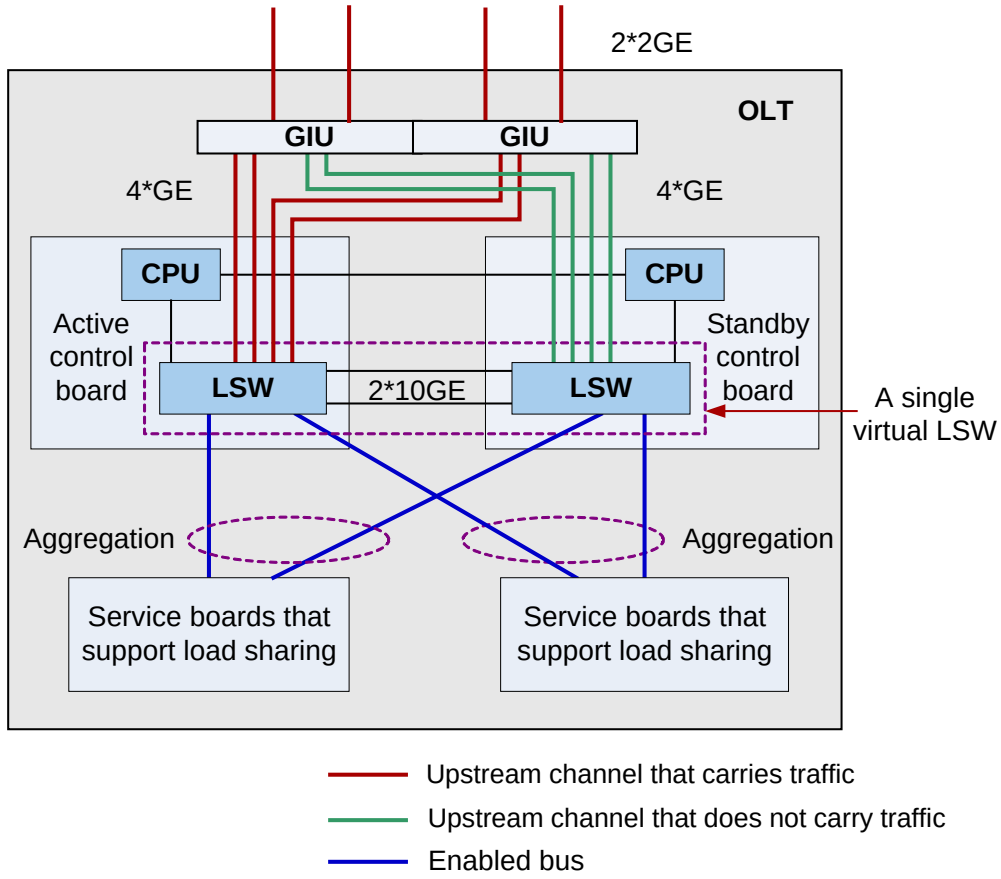
When the active control board functions properly, the buses between service boards and the active control board are enabled, but the buses between service boards and the standby control board are disabled. When detecting an switchover through hardware signals, service boards enable the buses between them and the new active control board (originally the standby control board) and disable the buses between them and the new standby control board (originally the active control board).

Load-sharing Mode

In load-sharing mode, both control boards forward data. Service boards that support load-sharing are connected to the active and standby control boards. The bandwidth between the service boards and control boards is doubled because the buses between the service boards and the control boards are enabled. The upstream bandwidth of the control board is the same as that of the GIU board. The SCUN board and GICK board are considered as an example. Figure 19-4 shows the principle of the load-sharing mode, which works as follows:

- On the forwarding plane,
  - Traffic between the service board and active and standby control boards is load balanced between the active and standby control boards. The LAN switch (LSW) chips on the active and standby control boards forward data at the same time and can be regarded as a single virtual LSW. If either control board fails or is removed, the system switching bandwidth is reduced by half.
  - The channel between the active control board and GIU board carries the upstream traffic. The upstream channel between the standby control board and the GIU board is in standby state and does not carry the upstream traffic.
- On the control plane, the active and standby control boards work in active/standby mode. The CPU on the active control board manages the system and controls data forwarding while the CPU on the standby control board is in the standby state and does not manage the system. The CPU of the active control board controls the LSW chip on the active control board. The CPU of the standby control board controls the LSW chip on the standby control board using configuration data that is synchronized from the CPU of the active control board.

Figure 19-4 Principle of load-sharing mode

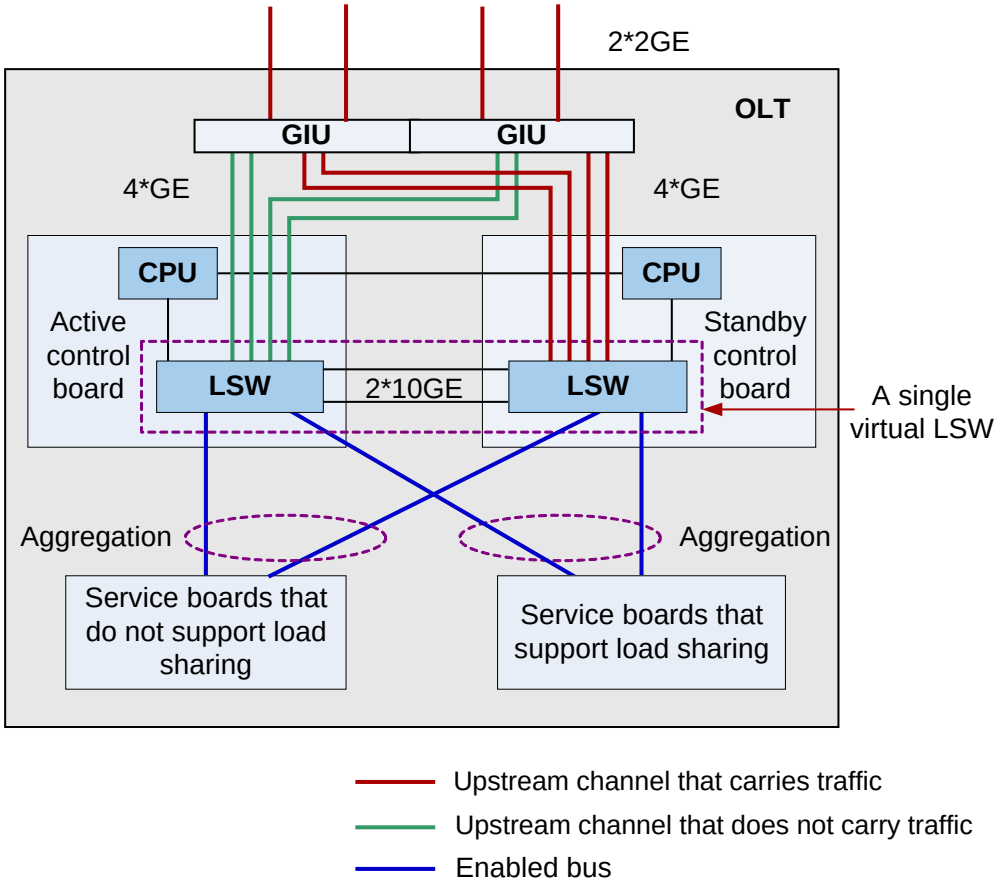


NOTE

In load-sharing mode, the upstream ports on both the active and standby control boards are in the working state, which is not presented in the figure.

After the active/standby switchover on the control plane, the upstream channel between the standby control board and GIU board carries the upstream traffic, as shown in Figure 19-5.

Figure 19-5 Principle of load-sharing mode (After the active/standby switchover)



Differences Between Active/Standby Mode and Load-sharing Mode

Device Switching Capability

The switching-capability differences between load-sharing mode and active/standby mode are as follows:

- In active/standby mode, only the active control board switches data. The standby control board is in the standby state and does not participate in data switching. Therefore, the system switching capability is determined by the switching capability of the active control board.
- In load-sharing mode, the system distributes service board data to the two control boards using a load-sharing algorithm, and both control boards switch data. Therefore, the system switching capability is determined by the switching capability of the active and standby control boards.

Upstream Port

The upstream port here refers to upstream ports on GIU boards and control boards.

- Upstream ports on GIU boards

Two GIU boards work concurrently regardless of the working mode (active/standby mode or load-sharing mode) of the control boards.

- Upstream ports on control boards

  When the control boards work in active/standby mode, only upstream ports on the active control board are working. When the control boards work in load-sharing mode, upstream ports on both of the control boards are working.

Service Boards

Service boards are connected to the active and standby control boards through buses. When the control boards work in load-sharing mode, the buses are enabled between the service boards that support load-sharing and the control boards so that the upstream bandwidth of the system is doubled.

Table 19-3 lists the working modes of service boards for different control board working modes.

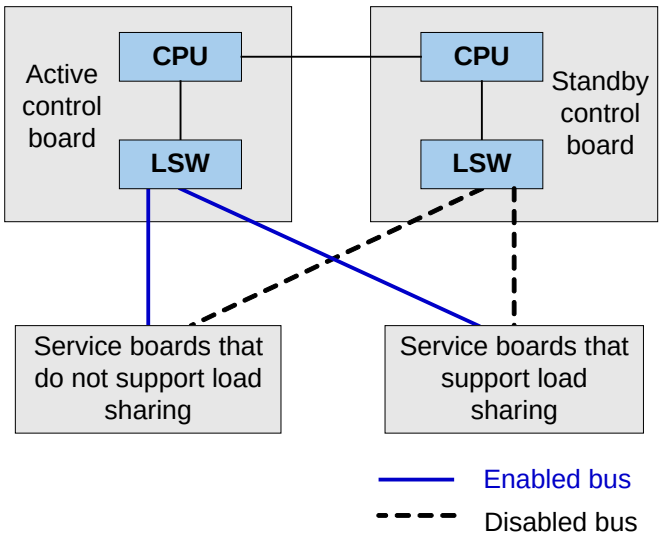
Table 19-3 Working modes of service boards

Service boards	Control Boards Working in Active/Standby Mode	Control Boards Working in Load-sharing Mode
Service boards that do not support load sharing	Active/standby mode	Active/standby mode
Service boards that support load sharing	Active/standby mode	Load-sharing mode

The working modes are as follows:

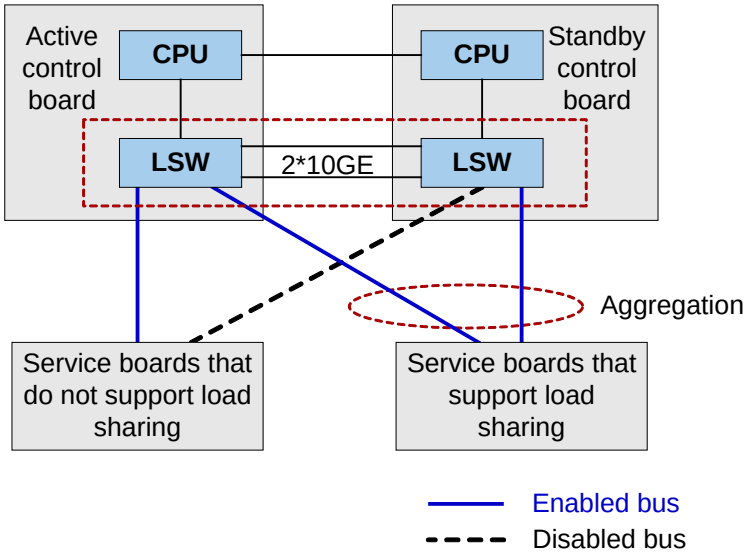
- Service boards that do not support load sharing always work in the active/standby mode, regardless of the working mode of the control boards. In other words, the buses between the service boards and the active control board are enabled, but the buses between the service boards and the standby control board are disabled, as shown in Figure 19-6 and Figure 19-7.
- Service boards that support load sharing work in the same mode as the control boards. The system automatically configures the working mode for service boards when the service boards start up.
  - If the control boards work in active/standby mode, the system sets the working mode to active/standby for the service boards. The default mode is the active/standby mode. The buses between the service boards and the active control board are enabled, but the buses between the service boards and the standby control board are disabled, as shown in Figure 19-6.

Figure 19-6 Connections between service boards and control boards in active/standby mode



- If the control boards work in load-sharing mode, the system sets the working mode to load sharing for the service boards. The buses between the service boards and the control boards are enabled and aggregated, as shown in Figure 19-7.

Figure 19-7 Connections between service boards and control boards in load-sharing mode



19.3 Ethernet Link Aggregation

Ethernet link aggregation is a process that aggregates two or more Ethernet ports of the same type to a logic port. It increases the link bandwidth without requiring a hardware upgrade and improves link reliability using the link backup mechanism.

19.3.1 What Is Ethernet Link Aggregation

As broadband services are widely used, carriers require higher bandwidth and reliability for Ethernet links.

Hardware upgrades can increase Ethernet link bandwidth but at high costs. In addition, hardware upgrades are less flexible than software upgrades. Ethernet link aggregation addresses these issues and supports the following functions:

- The maximum bandwidth of a link aggregation group (LAG) is equal to the total bandwidths of all links in a LAG. Hardware upgrades are not required and costs are curtailed.
- Traffic in a LAG is distributed to member links using the load sharing algorithm, implementing load sharing and improving link usage.
- Member links in a LAG dynamically back up each other. When one link is interrupted, a bakckup link immediately takes over.
- Link aggregation functions between interconnected devices only and is independent of the network topology.

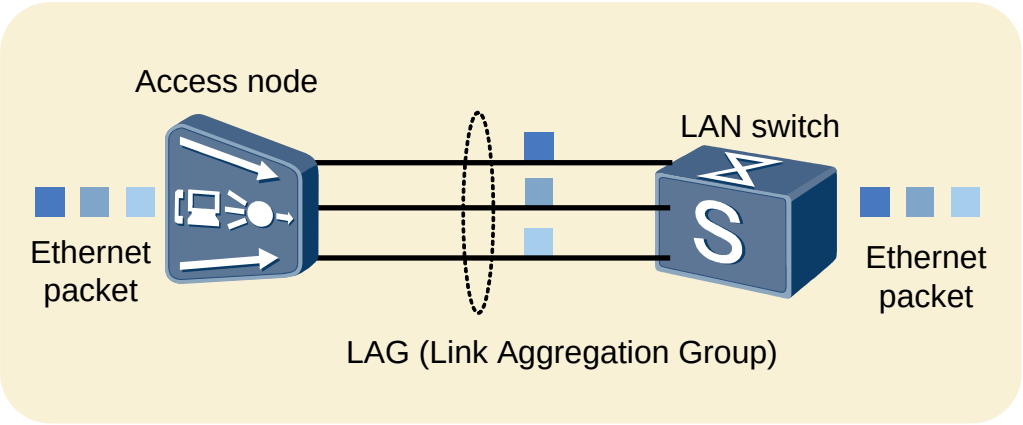
A logical link aggregating several physical links is called a LAG.

NOTE

Link aggregation is also called port aggregation because each link corresponds to two specific ports on an Ethernet network.

As shown in Figure 19-8, two devices are interconnected through several pairs of Ethernet ports. These ports are bound into a logical link, called a LAG.

Figure 19-8 Link aggregation



19.3.2 Basic Concepts of Ethernet Link Aggregation

LACP

Definition

LACP was developed based on IEEE 802.1AX. LACP dynamically aggregates and deaggregates Ethernet links using the following functions:

- Enables the data switching device to use a standard negotiation mode for link aggregation.

  After LACP is enabled, the access device aggregates multiple links according to its configuration and enables the aggregation link to transmit and receive data.

  Member ports in a LAG send LACP packets to notify the peer device of the information, including the system priority, system MAC address, port priority, port ID, and administrative key.

  After receiving the information, the peer device compares the information with the one saved on other ports to select ports that can be aggregated. The interconnected devices reach agreement on the ports that can transmit and receive data and determine which links carry traffic.

  The LACP packets are sent in any of the following modes:

  - Event-triggered transmission

    A change in the state of the local device or in the local configuration triggers the generation and transmission of a new LACP packet.

  - Periodic transmission

    When an aggregation link is stable, the system state is periodically exchanged to maintain the aggregation link.

- Maintains the aggregation link by periodically exchanging the system state when the link is stable.

  After aggregation links are generated, LACP maintains link status. When the aggregation condition changes, LACP automatically adjusts the link aggregation. The aggregation condition changes if one or more of the following occurs:

  - Physical port status changes.
  - Board status changes.
  - Results negotiated with the peer end are changed.

LACP Priority

LACP priority includes system priority and port priority. Priority setting allows negotiation of aggregation information between LAGs at two ends and real-time maintenance on link status.

Priority Type	Function	Description
System priority	Specifies the priority of a LAG on the device.	Determines which party in the interconnected devices is dominant in protocol negotiation. A device with a higher priority is dominant in LACP protocol negotiation, and the information provided by it prevails for such operations as selection of an active port. **NOTE** If interconnected devices have the same system priority, compare the system MAC addresses. The device with a smaller MAC address is dominant in LACP protocol negotiation.
Port priority	Specifies the priority of a port in a LAG.	Specifies the active ports in a LAG that will carry services with preference. Active ports in a LAG can be adjusted based on

Priority Type	Function	Description
		port priorities. **NOTE** In a LAG, if two ports have the same priority, the port with smaller subrack ID, slot ID, and port ID (*frameid/slotid/portid*) is preferentially selected as the active one.

NOTE
System priority and port priority work together to determine which ports in a LAG are used to carry services with preference. System priority prevails over port priority.

Timeout Time for Exchanging LACP Packets

The access device periodically sends and receives LACP packets to prevent information loss. If the device does not receive any LACP packets within three LACP packet exchanging periods, the device determines that the port is faulty.

To ensure detection sensitivity, LACP defines long timeout and short timeout.

The device uses short time for exchanging LACP packets by default except that the peer device requires long timeout.

According to LACP, the short timeout time is 1s, and the long timeout time is 30s. Huawei access devices support the following timeout time:

- Short timeout: 1-10s
- Long timeout: 20-40s

LACP Applications

Huawei access devices use LACP to implement link aggregation. Specifically, after an LACP LAG is configured, the access device negotiates with its peer end about aggregation information. For details about LACP applications, see "LACP Aggregation" in LAG Type.

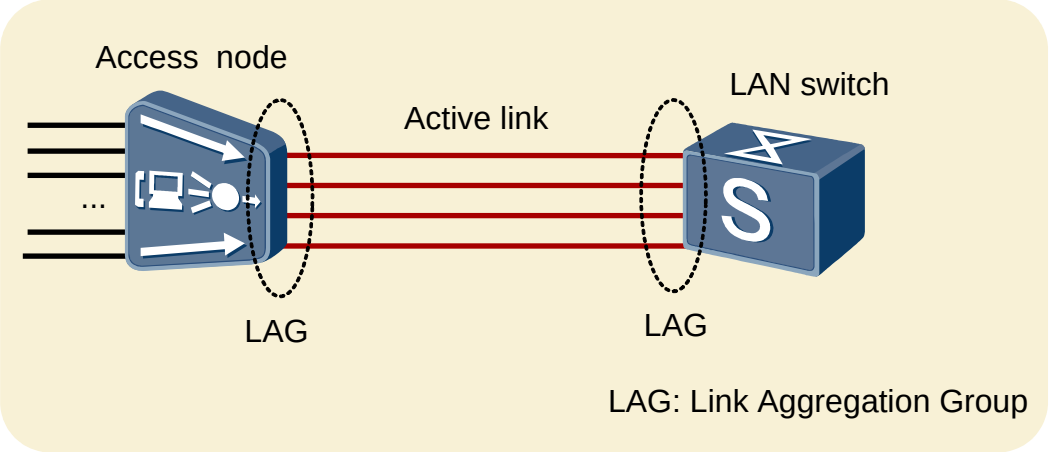
Load Bearing

The device supports load sharing and load non-sharing for a LAG. Both load sharing and load non-sharing increase link bandwidth and improve link reliability, but they achieve this in different ways.

Load Sharing Link Aggregation

Each member link in a load sharing LAG carries traffic. Member links in the LAG share the load, as shown in Figure 19-9.

Figure 19-9 Load Sharing link aggregation



To ensure packets carried on member links are correctly received on the peer device and ensure load balancing over member links in a LAG, the device allocates packets using hash algorithms based on:

- MAC addresses, including source MAC addresses, and source MAC addresses+destination MAC addresses
- IP addresses
- MPLS labels

NOTE
Applications vary depending on device capabilities.

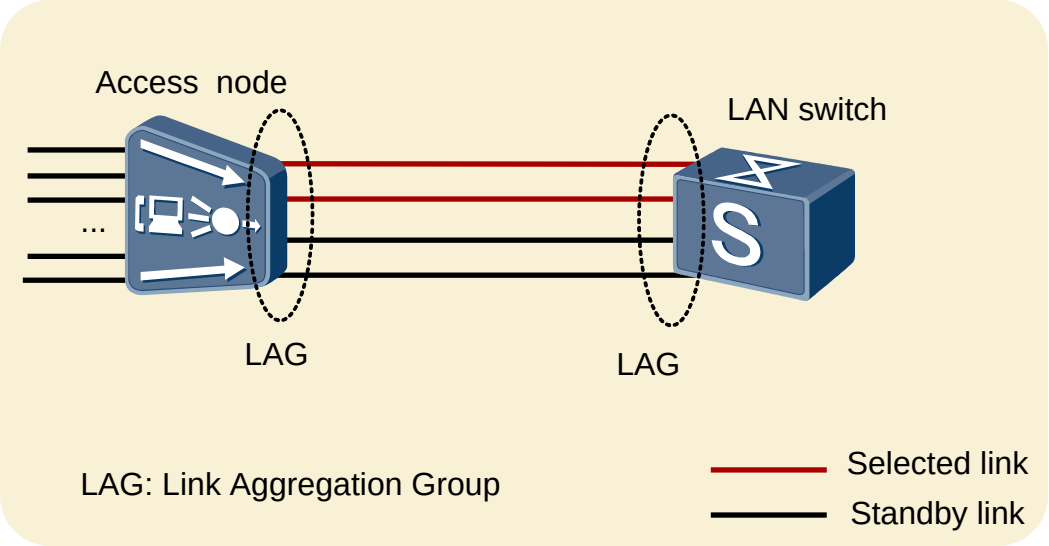
If member links in a LAG change or some member links become faulty, the device automatically reallocates traffic.

Load Non-Sharing Link Aggregation (Active/Standby Mode)

In a load non-sharing LAG, only some member links carry traffic and other links are in the standby state. Member links work in a hot standby mechanism, as shown in Figure 19-10.

NOTE
Load non-sharing link aggregation is implemented by configuring the maximum number of links in a LAG, and it applies to a scenario where the device is single-homed to the upstream device.

Figure 19-10 Load non-sharing link aggregation (active/standby mode)



When the selected link fails, the device selects a link from the standby links to take over.

LAG Type

The device supports manual aggregation and LACP aggregation for a LAG. Table 19-4 defines them and describes their characteristics. If both interconnected devices support LACP, LACP aggregation is recommended.

Table 19-4 LAG type and characteristics

LAG Type	Definition	Application Scenario	Impact on Services
Manual aggregation	LACP is not enabled on the device. The device determines whether to aggregate a port according to its physical status (up or down), working mode, and rate.	The device is interconnected with the device that does not support LACP.	Interconnected devices do not fully negotiate with each other to aggregate links. In this case, when links work in forced GE mode and the receive or transmit direction of a member link fails (for example, when the optical fiber connected to a receive or transmit Ethernet port is cut), or when the link is incorrectly connected, the service transmit end cannot detect the fault. As a result, data may be lost.
LACP aggregation	LACP is enabled on the device. By running LACP, interconnected devices have the same aggregation information, including physical port status,	The device is interconnected with the device that supports LACP.	None

LAG Type	Definition	Application Scenario	Impact on Services
	working mode, rate, and LACP priority. An LACP LAG has more accurate and effective control over link aggregation than a manual LAG.		

Table 19-5 lists the relationship between LAG type and load sharing mode.

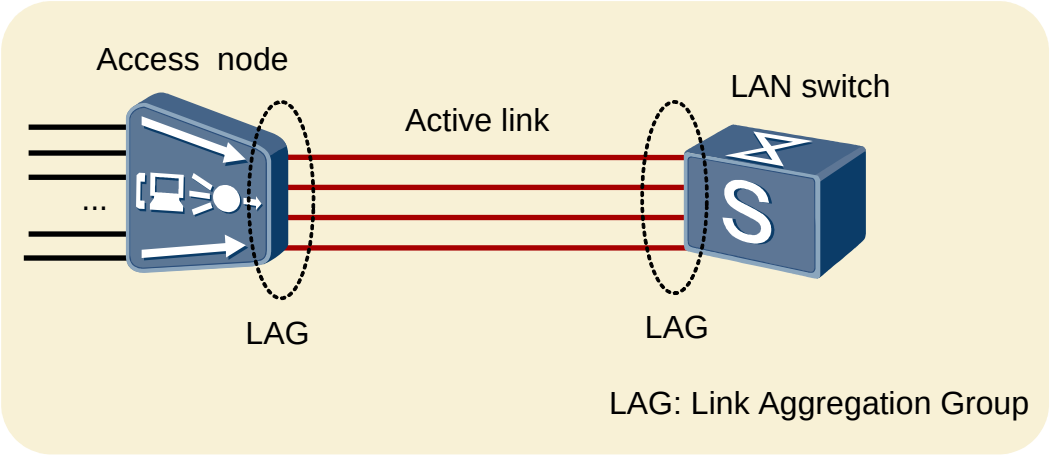
Table 19-5 Relationship between LAG type and load sharing mode

LAG Type	Load Sharing	Load Non-Sharing (Active/Standby Mode)
Manual aggregation	Supported	N/A
LACP aggregation	Supported	Supported

Manual Aggregation

Figure 19-9 shows manual aggregation. In this type, load is shared on all member links, and there is no standby link.

Figure 19-11 Load Sharing link aggregation



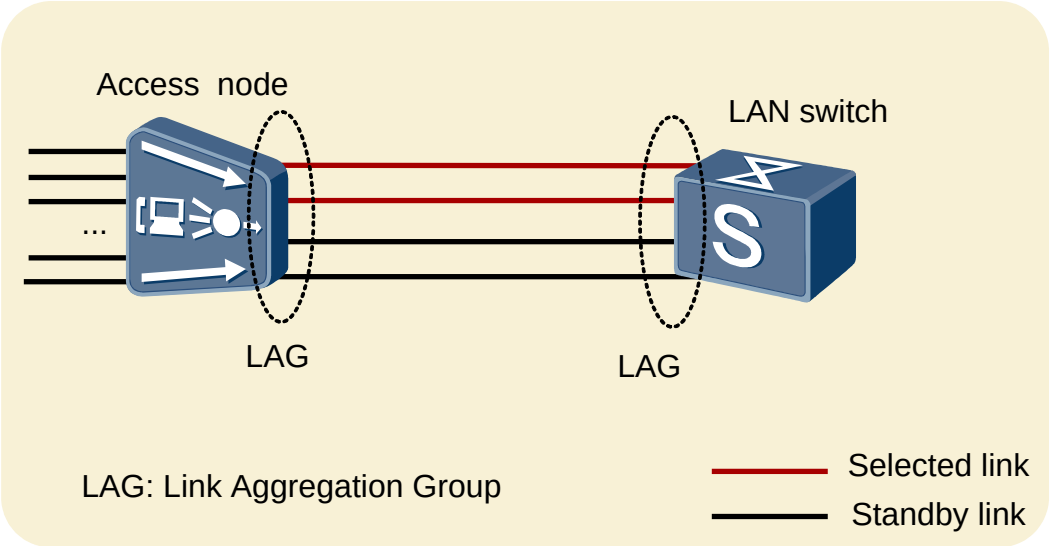
LACP Aggregation

LACP aggregation has the advantages of manual aggregation and LACP. In LACP aggregation, LACP is enabled only on the specified port, facilitating network topology control.

The network topology of LACP aggregation in load sharing is the same as that of manual aggregation, as shown in Figure 19-9.

Figure 19-10 shows LACP aggregation in load non-sharing.

Figure 19-12 Load non-sharing link aggregation (active/standby mode)



Member links in an LACP LAG are either selected links or standby links.

- A selected port is a working port that carries traffic.
- A standby port does not carry traffic.

Not all the member ports in an LACP LAG are working at the same time. The selected or standby state of a port changes with the change in system running or external environment.

- The port status change results in port status changes at the LACP layer. For example, if a port is faulty in a LAG, its state is switched to standby at the LACP layer.
- The LACPDU exchange may result in port status changes at the LACP layer. For example, a device may change the status of its ports after it receives an LACPDU from its peer device.

NOTE
For details about the frame structure of LACPDUs, see IEEE 802.1AX-2008.

Port Type

Member ports in a LAG are classified into primary ports and secondary ports.

Table 19-6 Port type and characteristics

Type	Definition	Port Characteristics	
Primary port	Logical port representing a LAG in service configuration, used in service	Similarity • When creating a LAG, you must specify both the	Difference • The primary port represents the LAG to participate in service configuration whereas secondary

Type	Definition	Port Characteristics	
	configuration and query.	primary and secondary ports. • The primary and secondary ports are defined logically.	ports cannot participate in service configuration.
Secondary port	All member ports except for the primary port.		• A LAG has only one primary port but can have several secondary ports. • The primary port can quit its affiliated LAG only after the LAG is deleted. A secondary port can be added to or deleted from a LAG. • After a LAG is deleted, its services are still carried by the primary port.

NOTE

- An Ethernet port can be added to only one LAG.
- If the primary port is faulty, secondary ports work. However, service-related operations can only be performed on the primary port.
- The device supports a LAG containing only one port. When the LAG requires expansion, you only need to add ports to the LAG but do not need to modify services.

19.3.3 LACP Aggregation Implementation Principles

In Link Aggregation Control Protocol (LACP) aggregation, a LAG is created, switched over, or switched back using LACP.

NOTE

A manual LAG does not use LACP. Therefore, implementation principles of a manual LAG are not described in the remainder of this document. For details, see IEEE 802.1AX.

LAG Setup Process

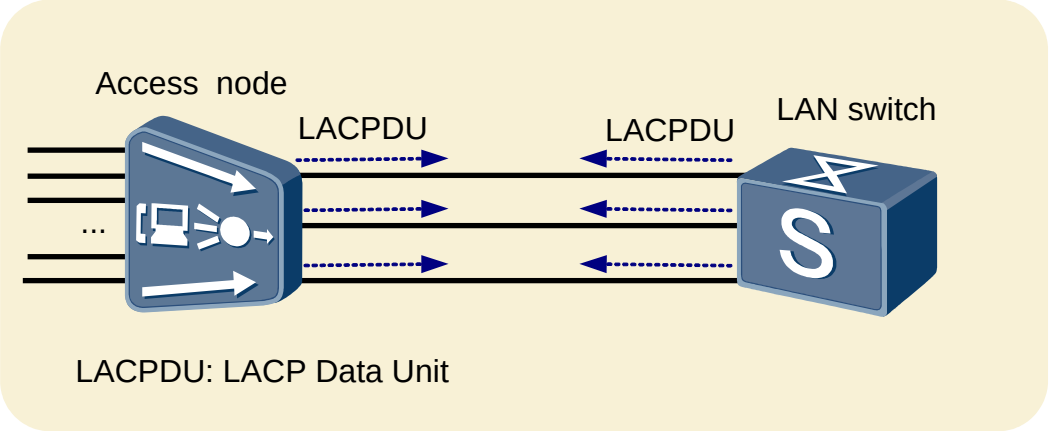
The process of setting up an LACP LAG is as follows:

1. Interconnected devices send LACPDUs to each other.
2. Interconnected devices determine the actor (indicating the party that is dominant in negotiation) based on LACP system priorities and system IDs.
3. Interconnected devices determine active ports (ports carrying traffic) based on the port LACP priorities and port IDs of the Actor.

Exchanging LACPDUs Between Interconnected Devices

An LACP LAG is set up on interconnected devices, and member ports are added to the LAG. The member ports are enabled with LACP, enabling interconnected devices to exchange LACPDUs to each other, as shown in Figure 19-13.

Figure 19-13 Exchanging LACPDUs between interconnected devices of an LACP LAG

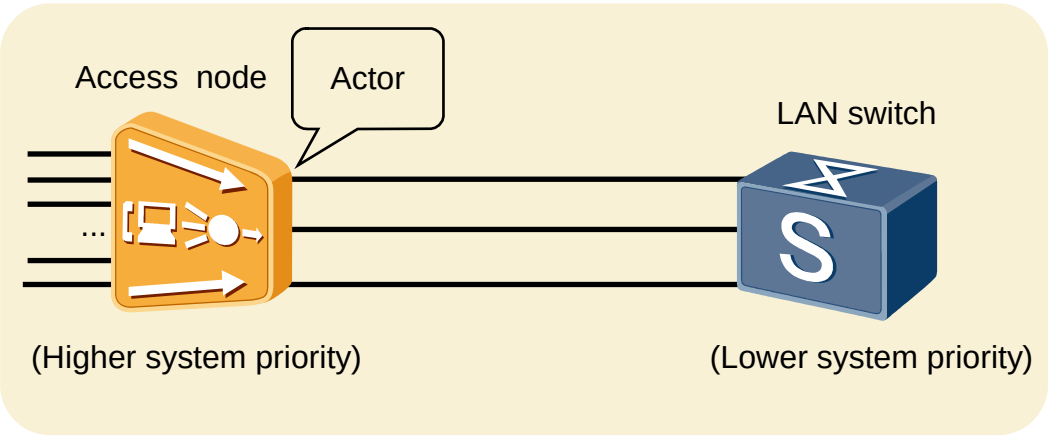


Determining the Actor

Interconnected devices receive LACPDUs from each other. Use the LAN switch shown in Figure 19-14 as an example to describe the process. After receiving LACPDUs from the access node, the LAN switch checks and records information about the access node and compares system priorities. If the system priority of the access node is higher than that of the LAN switch, the access node acts as the actor.

If the system priority of the access node is the same as that of the LAN switch, the party with a smaller MAC address in the system ID functions as the actor.

Figure 19-14 Determining the actor

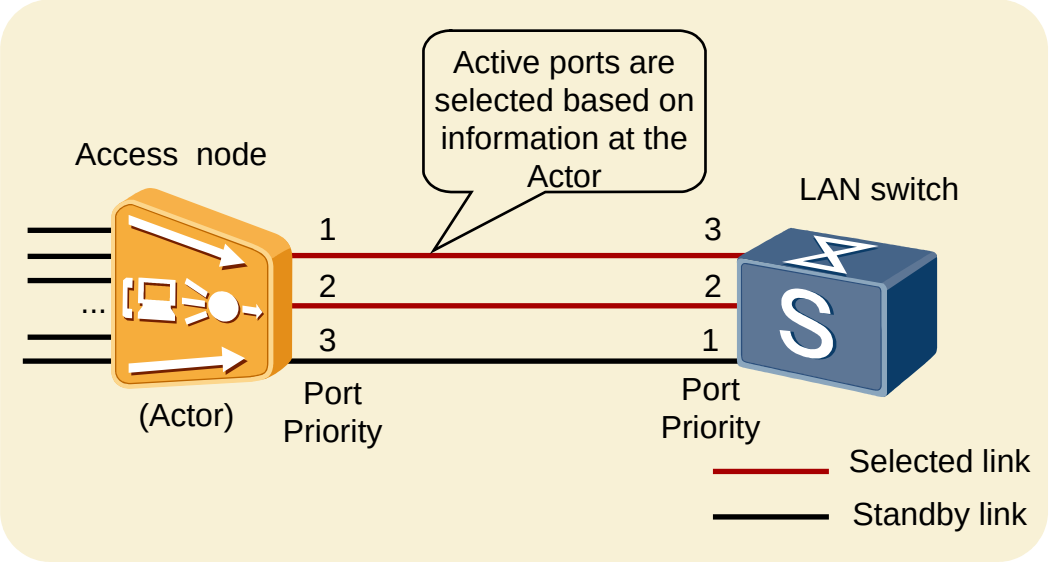


Selecting Active Ports

As shown in Figure 19-15, interconnected devices select active ports based on the priorities of ports on the actor after the actor is selected.

After the same active port is selected, selected links in the LAG are determined, and traffic is distributed in load sharing across the links.

Figure 19-15 Selecting active ports



LAG Link Switchover and Reversion

Switching Condition

In an LACP LAG, a link switchover is triggered if a device at one end detects one of the following events:

- LACP detects a link failure.
- Negotiated LACP status is changed.
- An active port becomes unavailable.
- The board is faulty.

Switchover Process

When any of the preceding trigger conditions is met, the link switchover is performed in one of the following processes:

For load non-sharing LAGs:

1. The faulty link is disabled.
2. The standby link with the highest priority is selected to replace the faulty selected link.
3. The standby link with the highest priority becomes the selected link and then forwards data.

For load sharing LAGs:

1. The faulty link is disabled.
2. Traffic is reallocated to member links using the load sharing algorithm.

Switchover Reversion Process

If an LACP LAG works in load sharing mode and the active port before switchover recovers, the original faulty link is enabled, and traffic is reallocated using the load sharing algorithm.

If an LACP LAG works in load non-sharing mode, you can set the link revertive mode to revertive or non-revertive.

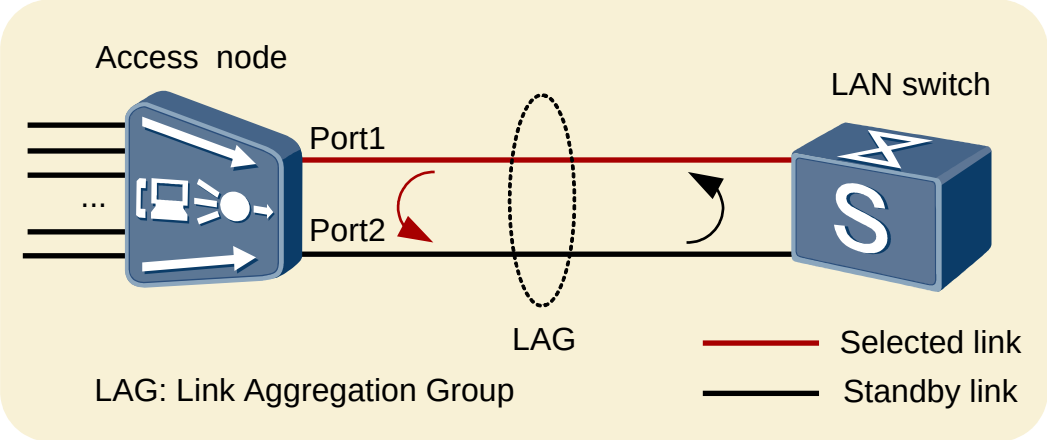
- When an LACP load non-sharing LAG is set to revertive, services are switched back to the link of a higher priority after this link recovers, as shown in Figure 19-16.
- When an LACP load non-sharing LAG is set to non-revertive, services are not switched back to the selected link after this link recovers, but are still transmitted on the current link.

As shown in Figure 19-16, ports 1 and 2 are member ports in an LACP LAG, port 1 is the active port (working in the selected link), and port 2 is the inactive port (working in the standby link).

After port 1 fails, port 2 takes over as the active port.

After port 1 recovers, port 1 works as an active port, and port 2 still works as an inactive port.

Figure 19-16 Reversing a switchover (LACP LAG)



19.3.4 Ethernet Link Aggregation Network Applications

SCU, ETHB, SPUA, and GIU boards can be used for common upstream aggregation.

Generally, ETHB, SPUA, and OPGD boards are recommended for downstream cascaded aggregation.

When an access device provides upstream ports, the upstream ports are generally configured in link aggregation groups (LAGs) for the protection purpose. The table below lists the recommended application of link aggregation.

Link Aggregation Mode	Application Scenario	Required Software Version
Manual aggregation	Single homing	—

Link Aggregation Mode	Application Scenario	Required Software Version
Link aggregation control protocol (LACP) aggregation in load sharing mode	Single homing	MA5600T/MA5603T/MA5608T V800R007C00 or later
LACP aggregation in primary/secondary mode	Single homing or dual homing	MA5600T/MA5603T/MA5608T V800R006C02 or later

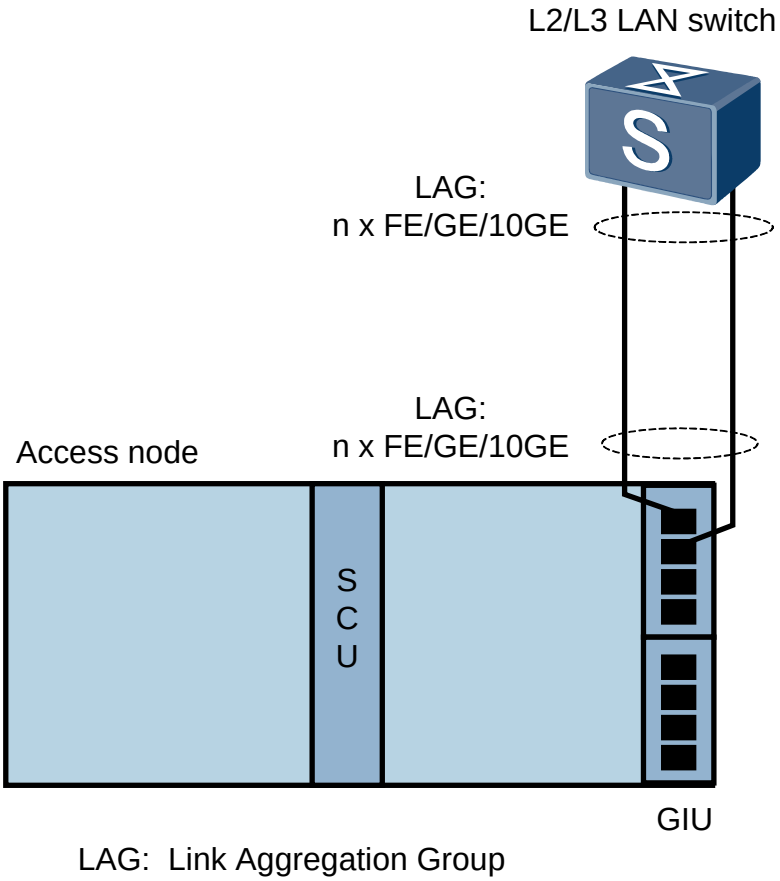
Upstream Transmission of Intra-Board Link Aggregation

Intra-board link aggregation configures two or more ports on a board into a link aggregation group (LAG).

Figure 19-17 shows an example of intra-board link aggregation: A GIU on the access device is used for upstream transmission and the access device is interconnected with another device. One LAG is configured on the GIU board of the access device and the other LAG is configured on one board of the interconnected device.

This type of network topology increases bandwidths (through load sharing) and protects links. However, if the board involved in LAG configuration fails, no protection is available.

Figure 19-17 Network topology of upstream transmission of intra-board link aggregation

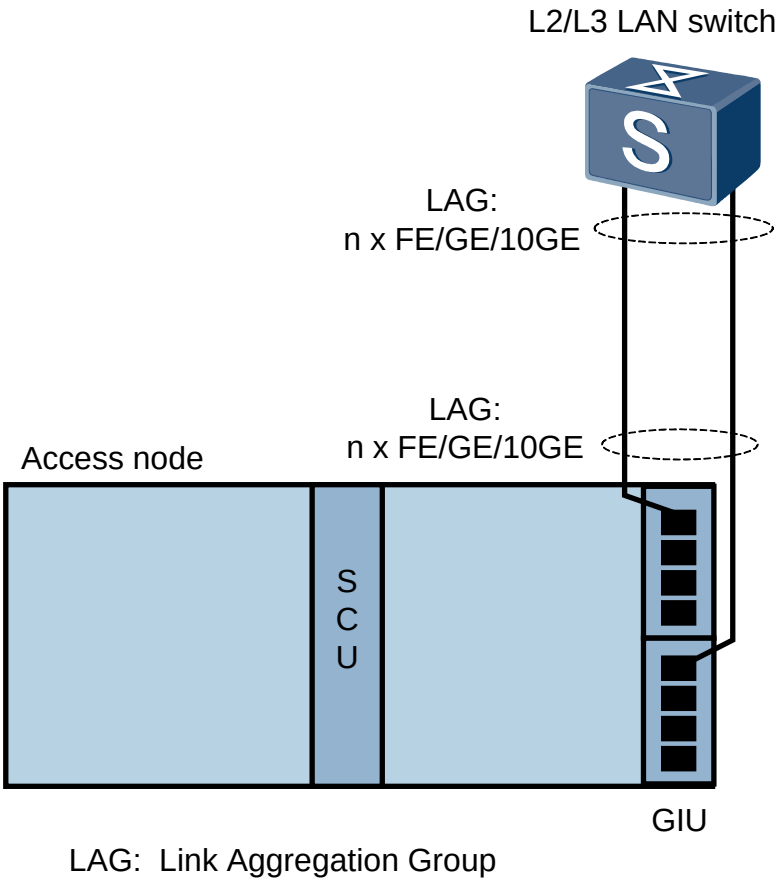


Upstream Transmission of Inter-Board Link Aggregation (Single Homing)

Inter-board link aggregation configures one or more ports on two boards into a LAG. Inter-board link aggregation provides protection for boards configured with aggregated links.

Figure 19-18 shows an example of inter-board link aggregation: Two GIUs on the access device are used for upstream transmission and the access device is interconnected with another device. One LAG is configured on the two GIU boards of the access device and the other LAG is configured on two boards of the interconnected device. This type of network topology increases bandwidths (through load sharing) and protects links. Protection is provided if the board involved in LAG configuration fails.

Figure 19-18 Network topology of upstream transmission of inter-board link aggregation (single homing)



Upstream Transmission of Inter-Board Link Aggregation (Dual Homing)

Dual homing in inter-board link aggregation refers to the following scenario: Primary and secondary upper-layer devices are configured; LAGs are configured on these two devices for interconnecting with the device. This type of link aggregation provides protection for upper-layer devices between LAGs. If the active LAG is functioning inappropriately, the bandwidth and priority of the active LAG are lower than those of the standby LAG. As a result, protection switchover is triggered.

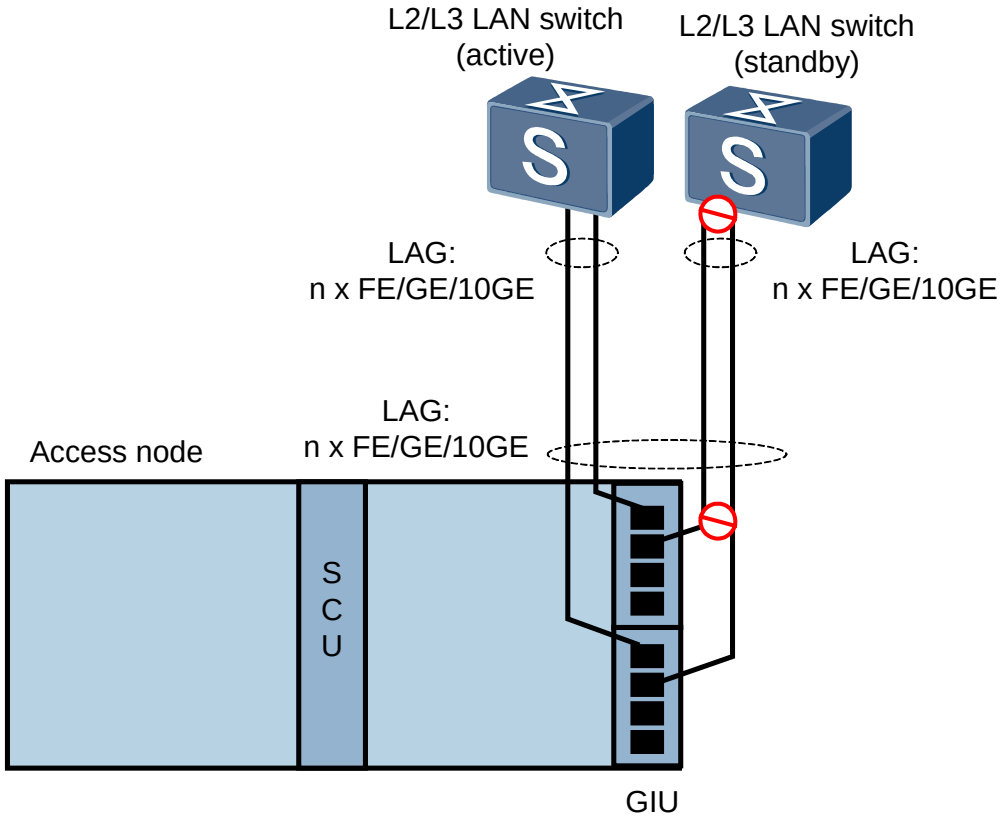
In upstream transmission of inter-board link aggregation scenarios, the access device generally functions as the master-end device to determine the primary and secondary roles in the LAG. To determine the primary device, the access device adheres to the following rules:

- The access device uses the received peer end's information carried in the LACP packet to determine the dual homing network. Then, the access device automatically calculates the number of **LinkUp** ports on the two upper-layer devices, and selects the device that has the greater number of **LinkUp** ports as the primary device. The other upper-layer device is the secondary device.
- If the two upper-layer devices provide the same number of **LinkUp** ports, the rules vary depending on the LACP priority preemption mode configured on the access device.

  - If LACP works in priority preemption mode, the access device preferentially selects the peer end device connected to the local port with the higher priority as the primary device.
  - If LACP works in non-priority preemption mode, the access device preferentially selects the peer end device connected to the port that is currently in forwarding state as the primary device.
- If the link forwarding traffic disconnects, the access device uses the preceding rule to select the primary device.

Figure 19-19 shows another example of inter-board link aggregation: Two GIUs on the access device are used for upstream transmission and the access device is interconnected with two devices. One LAG is configured on different boards of each of the interconnected devices.

Figure 19-19 Network topology of upstream transmission of inter-board link aggregation (dual homing)



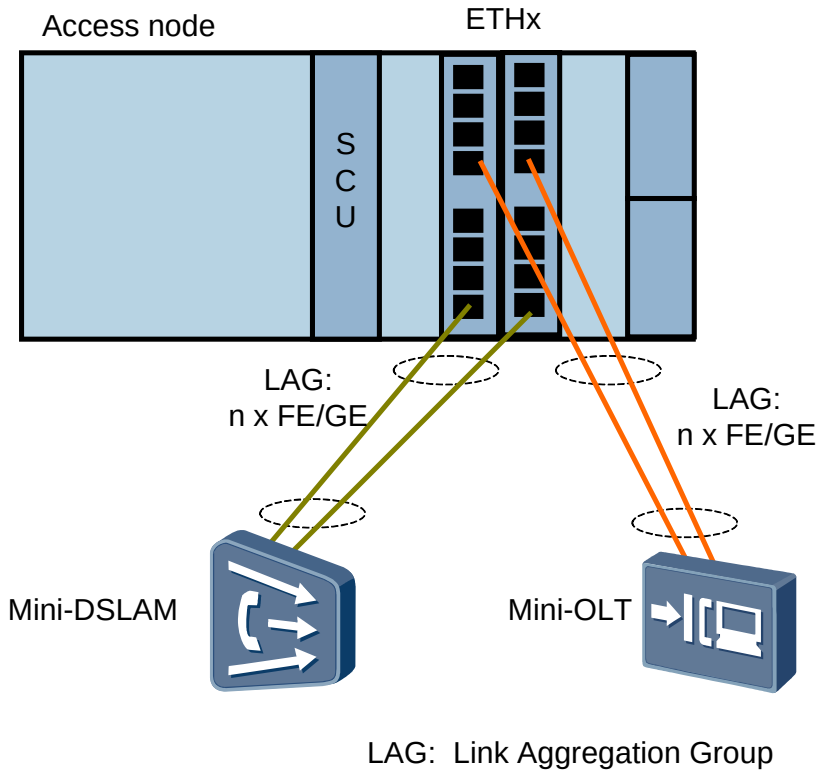
In this scenario, the access device determines the active and standby LAGs after its LACP system priority is configured higher than that of the peer devices. Alternatively, the upper-layer devices themselves can determine the active LAG if they support LACP aggregation groups between devices.

Inter-Board Aggregation Cascading

In Ethernet cascading, the access device cascades and converges the mini-DSLAM or mini-OLT close to the user. In this case, bandwidth and reliability of links are improved and resources of the upstream port on the CO equipment are reduced.

Figure 19-20 shows an example of inter-board aggregation cascading: Two ETH boards in inter-board link aggregation cascade the lower-layer devices. LAGs are configured on the interconnected devices.

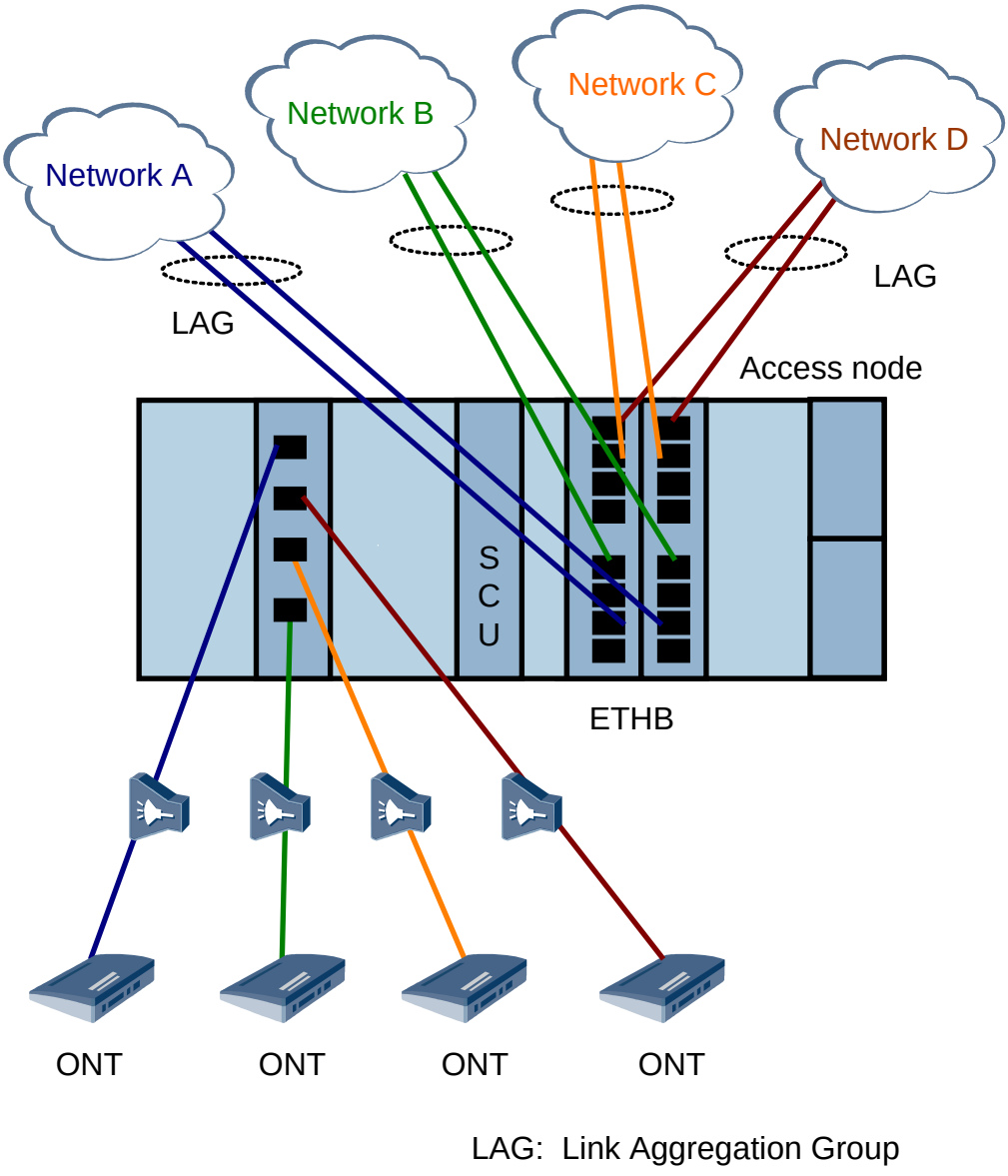
Figure 19-20 Network topology of inter-board aggregation cascading



Upstream Transmission of Inter-Board Link Aggregation (Multiple Services)

Carriers may provide various services for users such as Internet access service for common users, Internet private line service for enterprise users, VPN service for enterprise users, and IPTV service for common users. Different users require their own upstream ports for connecting to the MAN and require assured high bandwidth and reliable links. To meet user's requirements, use upstream transmission of inter-board link aggregation. Figure 19-21 shows an example of inter-board link aggregation for multiple services, in which two ETHB boards providing four groups of upstream ports are used for upstream transmission. These four groups are configured with four LAGs each and connected to different upper-layer networks.

Figure 19-21 Network topology of upstream transmission of inter-board link aggregation (multiple services)



19.3.5 Configuring Ethernet Link Aggregation

Configure Ethernet link aggregation to increase link bandwidth and improve link reliability, without performing a hardware upgrade.

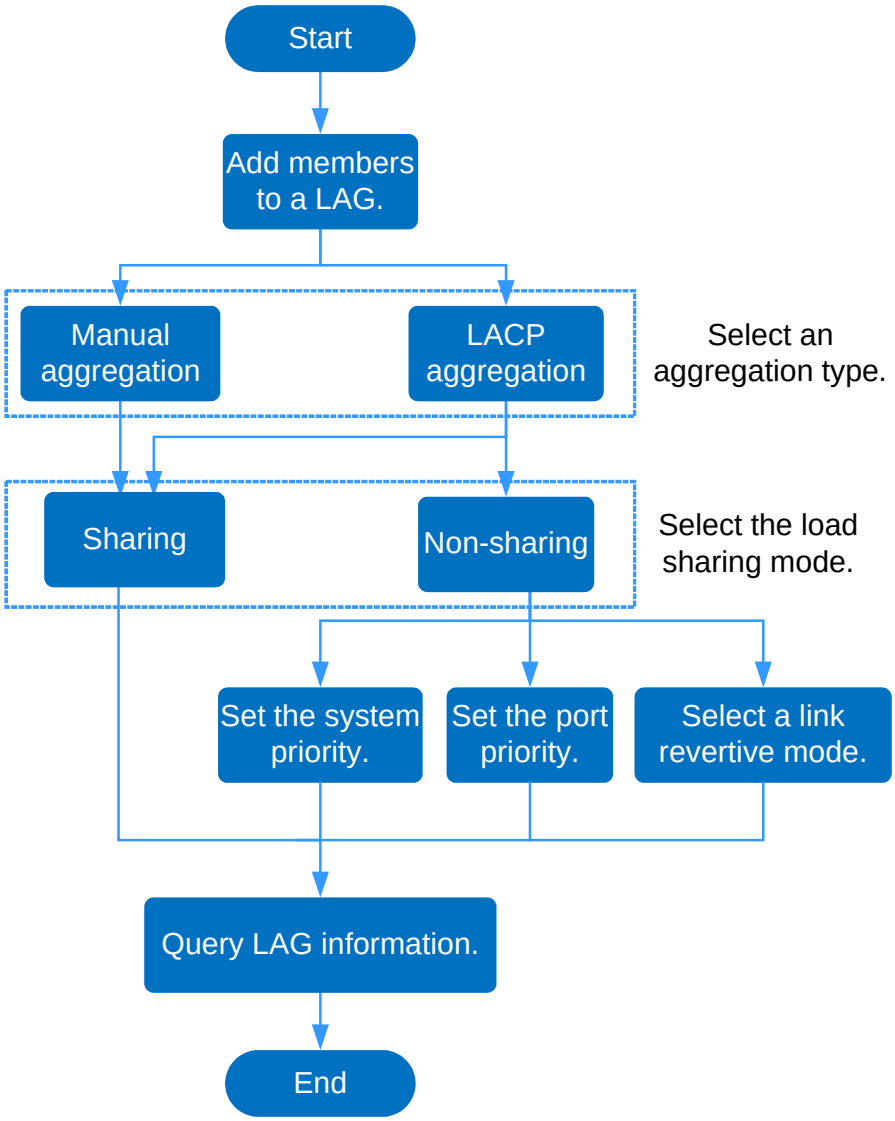
Prerequisites

- Interconnected devices, hardware, and port attributes must support LAGs.
- The two aggregated ports do not have static MAC addresses. You can run the **display mac-address** command to query whether an aggregated port has static MAC address.

Context

Figure 19-22 shows the flowchart for configuring a LAG.

Figure 19-22 Configuration flowchart



Procedure

(Mandatory) Create a LAG and select the aggregation type.

Run the **link-aggregation** command to add multiple upstream Ethernet ports to the same LAG to implement protection and load sharing between ports. The port with the smallest port ID is the master port.

NOTE

If the device is interconnected with the device that supports LACP, static aggregation is recommended. If the device is interconnected with the device that does not support LACP, only manual aggregation can be used.

Step 1 (Optional) Add a LAG member port.

Perform this step when the LAG bandwidth or link reliability needs to be improved further. Run the **link-aggregation add-member** command to add an Ethernet port to an existing LAG to increase the LAG bandwidth and improve the link reliability.

NOTE
If the port to be added to or deleted from a LAG is connected to the peer device, run the **shutdown(Ethernet)** command to deactivate the Ethernet port or remove the optical fiber to prevent loops.

Step 2 (Optional) Select the load carrying type.

NOTE
If the load sharing type is not configured, a LAG works in load sharing mode by default.

This step is required only when a static LAG is configured. Configuring the maximum selected links in a LAG implements traffic allocation in load non-sharing mode. For example, M+N links are configured in a LAG. Then, run the **link-aggregation max-link-number** command to specify N selected links. The remaining M links are standby ones. If a selected link is disconnected, a standby link automatically changes to the selected one.

Step 3 (Optional) Set the system priority and port priority.

This step is required only when a static LAG is configured.

- LACP system priority: If the access device is dual homed to two convergence devices, the access device determines the selected and standby LAGs. Run the **lacp priority system** command to set the LACP system priority of the access device to be higher than that of the peer device.
- LACP port priority: LACP port priority must be used together with the maximum number of links. If a port is required preferentially for carrying services, set its priority higher. Run the **lacp priority port** command to change the link priority so that the standby link and the selected link can be switched over.

Step 4 (Optional) Selected the link revertive mode.

This step is required only when a static LAG in load non-sharing mode is configured. Run the **lacp preempt** command to set whether traffic is switched back to the original link if the link failure is rectified.

Step 5 (Optional) Query LAG information.

Run the **display link-aggregation** command to query the LAG information, including primary port, number of links, aggregation type (manual or static), and maximum number of links.

----End

Example

Assume the following configurations: The MA5600T/MA5603T/MA5608T transmits services upstream using the GIU board, upstream ports 0/19/0 and 0/19/1 on the same GIU board are configured in an upstream port LAG, packets are distributed to the LAG member ports according to the source MAC address, and the working mode is LACP static aggregation. To perform these configurations, run the following commands:

NOTE

The network topology is shown in "Upstream Transmission of Intra-Board Link Aggregation" in 19.3.4 Ethernet Link Aggregation Network Applications.

```
huawei(config)#link-aggregation 0/19 0-1 ingress workmode lacp-static
huawei(config)#display link-aggregation all
  ------------------------------------------------------------------------
  Master port  Link aggregation mode  Port NUM  Work mode  Max link number
  ------------------------------------------------------------------------
  0/19/0      ingress                   2  lacp-static            -
  ------------------------------------------------------------------------
  Total: 1 link aggregation(s)
```

Assume the following configurations: The MA5600T/MA5603T/MA5608T transmits services upstream using the GIU board, upstream ports 0/19/0 and 0/20/0 on the active and standby GIU control boards are configured in an inter-board LAG, packets are distributed to the LAG member ports according to the source MAC address and destination MAC address, and the working mode is LACP static aggregation. To perform these configurations, run the following commands:

NOTE

The network topology is shown in "Upstream Transmission of Inter-Board Link Aggregation (Single Homing)" in 19.3.4 Ethernet Link Aggregation Network Applications of the Feature Description.

```
huawei(config)#link-aggregation 0/19 0 0/20/0 0 egress-ingress workmode lacp-static


huawei(config)#display link-aggregation all
  ------------------------------------------------------------------------
  Master port  Link aggregation mode  Port NUM  Work mode  Max link number
  ------------------------------------------------------------------------
  0/19/0      egress-ingress             2  lacp-static            -
  ------------------------------------------------------------------------
  Total: 1 link aggregation(s)
```

Assume the following configurations: The MA5600T/MA5603T/MA5608T is configured with only one control board, the SCUN control board and the GIU board are configured in an inter-board LAG, packets are distributed to the LAG member ports according to the source MAC address, and the working mode is LACP static aggregation. To perform these configurations, run the following commands:

```
huawei(config)#link-aggregation 0/9 0-3 0/19 0-1 ingress workmode lacp-static

huawei(config)#display link-aggregation all
  ------------------------------------------------------------------------
  Master port  Link aggregation mode  Port NUM  Work mode  Max link number
  ------------------------------------------------------------------------
  0/9 /0      ingress                   6  lacp-static            -
  ------------------------------------------------------------------------
  Total: 1 link aggregation(s)
```

19.3.6 Ethernet Link Aggregation Standards and Protocols Compliance

Ethernet link aggregation complies with IEEE 802.1AX-2008, IEEE standard for local and metropolitan area networks (link aggregation).

19.4 Ethernet Port Protection Group

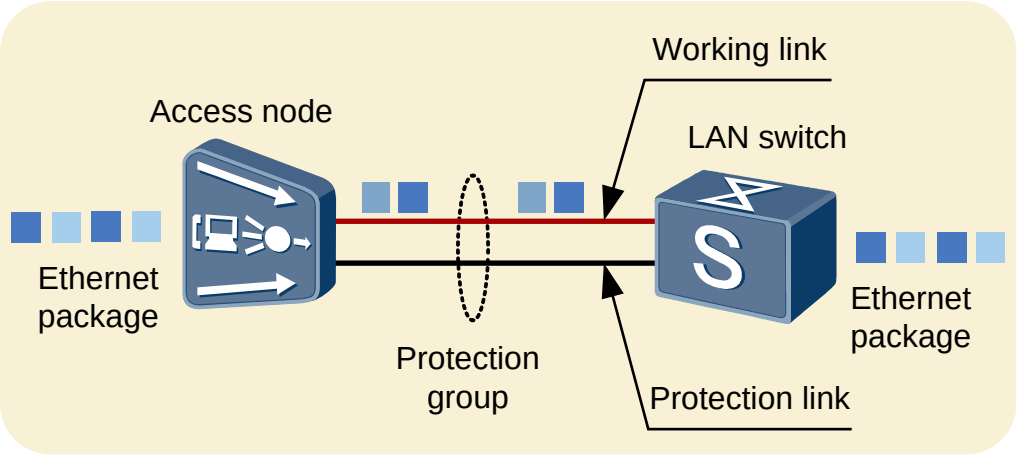
An Ethernet port protection group binds two or more Ethernet ports together. The working port carries services, while the protection port provides backup and does not carry services. If the working port is faulty, the services carried by it are switched to the protection port, automatically or manually as required. Ethernet port protection ensures uninterrupted service transmission and improves reliability of links.

19.4.1 Introduction to Protection Group of Ethernet Ports

Access network users, especially corporations and large residential communities, are posing stringent requirements on network reliability. If the upstream links of access devices are not protected, services carried over these upstream links will be interrupted if the links break or the upper-layer devices are malfunctioning. For the access devices that provide services to tens of thousands of users, carriers usually demand protection for the upstream ports. Specifically, services can be switched between links automatically or manually, depending on your requirements, for example, when the upstream link is broken. Such a protection group ensures that an upstream link is always available, improving reliability.

Figure 19-23 shows an example of an Ethernet port protection group between two interconnected devices. In this protection group, two Ethernet ports are bound. One Ethernet port protection group has two members. One member carries services over the working link, and the other member provides backup over the protection link.

Figure 19-23 Example of Ethernet port protection



An active/standby switchover implements protection group for upstream Ethernet ports on the control boards. According to the number of upstream Ethernet ports, members in a protection group an be ports, LAGs, or boards. If a working member on the active control board is faulty or a forced active/standby switchover is triggered manually, an active/standby switchover is triggered. After the switchover, the system switches services to protected members on the standby control boards.

Access devices support two types of Ethernet port protection groups: Portstate and Timedelay.

- A Portstate Ethernet port protection group switches services based on port status. The protection port is always enabled in this protection group, even when the working port is functioning properly. If the working port is faulty, the protection group immediately switches services to the protection port if the protection port is functioning properly.

- A Timedelay Ethernet port protection group also switches services based on port status, but the protection port is disabled when the working port is functioning properly. If the working port is faulty, the protection port remains enabled for a predefined period of time. Services are switched to the protection port if this port is functioning properly during this period. If the protection port is malfunctioning during this period, the Timedelay protection group retries the working port while it keeps the protection port disabled. The protection group will repeat the preceding process until one port restores the normal state, and then switch services to it.

Choose a protection type that suits your application scenario based on the types of interconnected devices and boards.

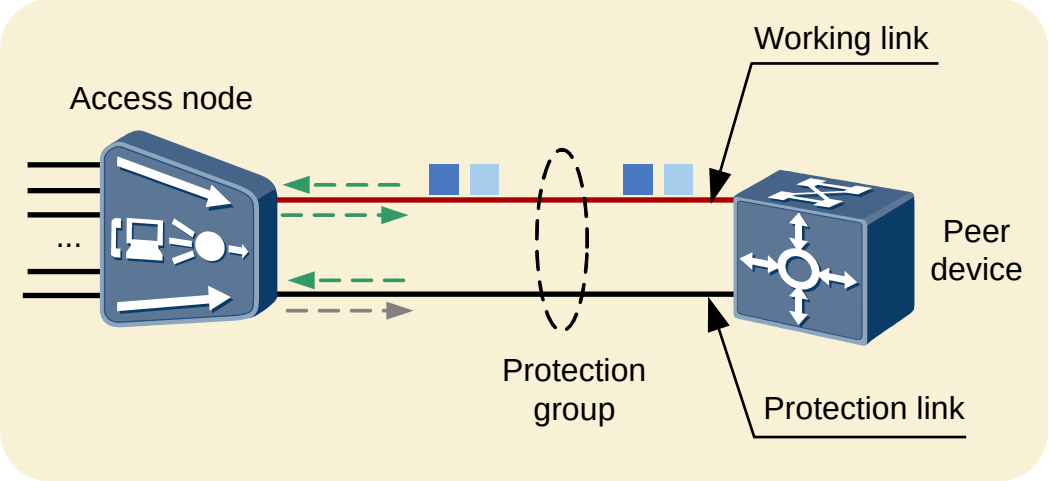
19.4.2 Principle of Portstate Protection

Establishment of a Portstate Protection Group

A Portstate protection group is established after it is configured, and the interconnected devices implement protection without the need of negotiation.

In the protection group, the working port is enabled to transceive data and a service is transmitted over the link carried by the port; the protection port is enabled only to receive data and no service is transmitted over the link carried by it, as shown in Figure 19-24.

Figure 19-24 Data transceiving within a Portstate protection group



Protection Switching

A Portstate protection group undergoes protection switching if the working port is faulty. The specific conditions for protection switching vary according to the protection level. For details, see Table 19-7.

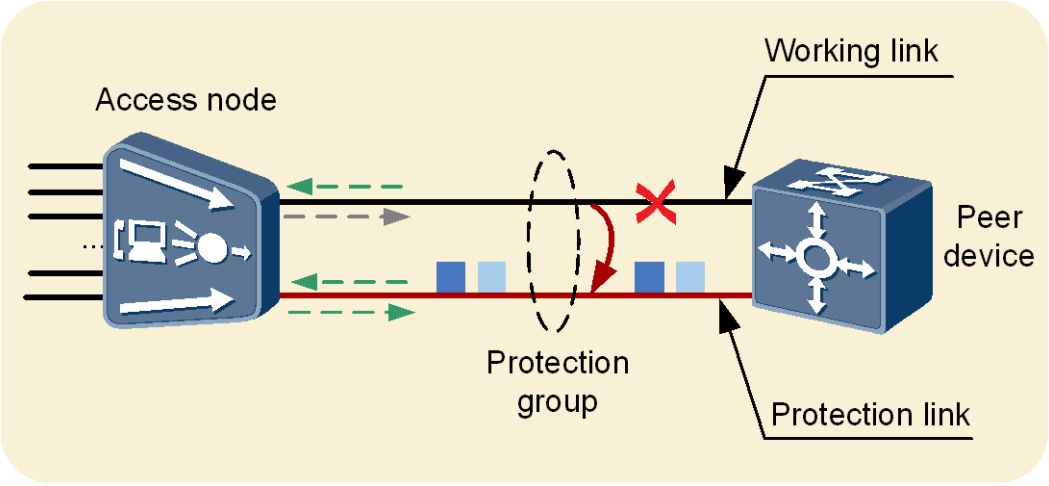
Table 19-7 Conditions for Portstate protection switching

Protection Level	Working Member - Protection Member	Switching Condition
Link level	One Ethernet port on the	• The working port is faulty.

Protection Level	Working Member - Protection Member	Switching Condition
	active control board - One Ethernet port on the standby control board	• Users manually start a forced protection switching.
LAG level	Ethernet LAG on the active control board - Ethernet LAG on the standby control board	• The working port is faulty and the LAG on the active control board has fewer properly-functioning ports than the LAG on the standby control board. • Users manually start a forced protection switching.
Board level (all ports on the board are protected)	Active control board - Standby control board	• The working port is faulty and the active control board has fewer properly-functioning ports than the standby control board. • Users manually start a forced protection switching.

If the working link is faulty, services are switched to the protection link to ensure uninterrupted forwarding, as shown in Figure 19-25.

Figure 19-25 Data transceiving within a Portstate protection group (after protection switching)



NOTE
The preceding figure illustrates the protection switching process for a link-level Portstate protection group. For an LAG-level or board-level Portstate protection group, LAGs or boards are considered as members.

Portstate protection switching is implemented as follows:

1. The access device disables the working port after detecting a working port fault. The working port is disabled for 90s, and during this period the access device notifies the interconnected device of the working link failure.

2. The interconnected device enables its protection port after detecting the working link failure.
3. The access device enables its properly-functioning protection port and switches services to the protection link.

In this protection switching process, the access device plays a "passive" role, because it starts protection switching after the interconnected device.

Protection Reversion

If the protection switching is triggered by a forced switching command or a working port failure, services will not be switched back to the original link even if the original working port restores the normal state. Instead, the services are still carried over the protection link.

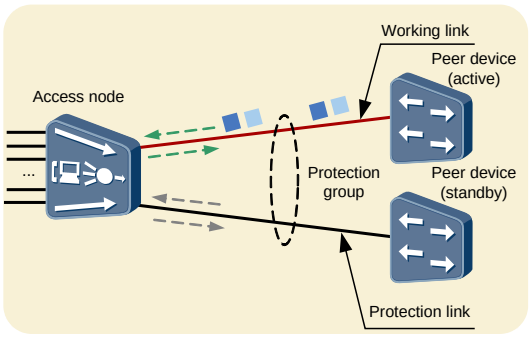
19.4.3 Principle of Timedelay Protection

Establishment of a Timedelay Protection Group

A Timedelay protection group is established after it is configured and the interconnected devices implement protection without the need of negotiation.

In the protection group, the working port is enabled to transceive data and a service is transmitted over the link carried by the port; the protection port is disabled and no service is transmitted over the link carried by it. In the Timedelay protection scenario, the access device can be either single-homed or dual-homed to the upper-layer devices. Figure 19-26 shows how data is transceived in the dual-homing scenario.

Figure 19-26 Data transceiving within a Timedelay protection group



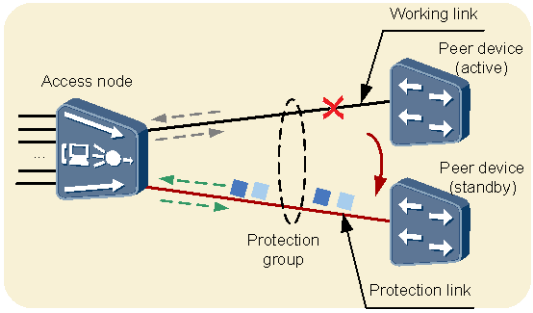
Protection Switching

A Timedelay protection group undergoes protection switching on any of the following conditions:

- The working port is faulty.
- Users manually start a forced protection switching.

If the working link is faulty, services are switched to the protection link to ensure uninterrupted forwarding, as shown in Figure 19-27.

Figure 19-27 Data transceiving within a Timedelay protection group (after protection switching)



Timedelay protection switching is implemented as follows:

1. The access device detects a working port failure and immediately enables its protection port.
2. The access device plays an "active" role and immediately switches services after detecting a working port failure. Services are then switched to the protection link.

NOTICE

The working and protection ports on peer devices must be enabled.

Protection Reversion

If the protection switching is triggered by a forced switching command or a working port failure, services will not be switched back to the original link even if the original working port restores the normal state. Instead, the services are still carried over the protection link.

19.4.4 Protection Group of Ethernet Ports Network Application

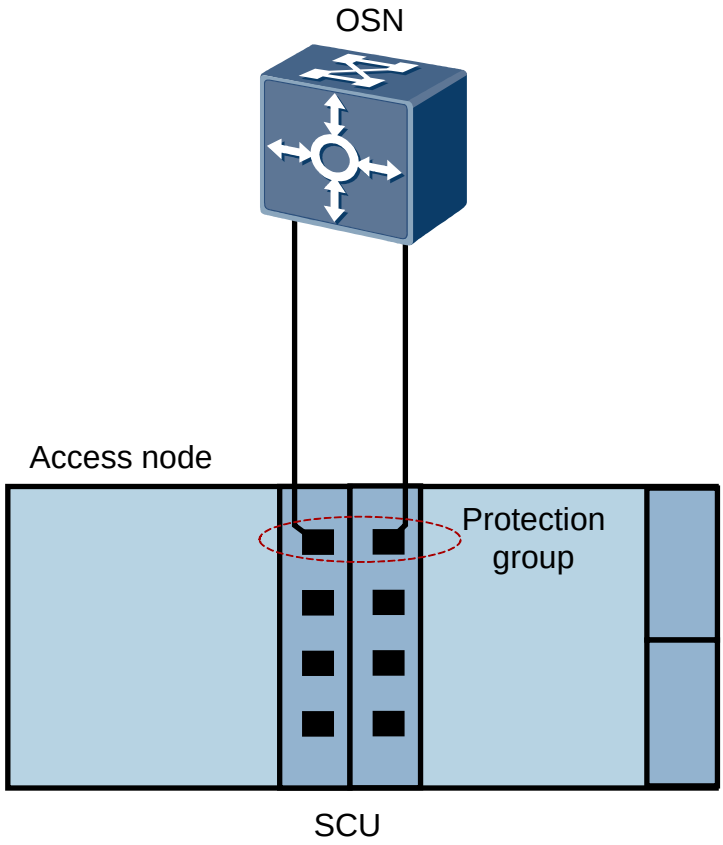
The system supports two types of Ethernet port protection groups:

- Portstate protection group: Applies to a network scenario in which the control board provides upstream ports, and protection is available at link level, LAG level, or board level. In the upstream direction, the access device is usually single-homed to a transmission device.
- Timedelay protection group: Applies to a network scenario in which the system control board or upstream board provides upstream ports, and protection is available only at link level. In the upstream direction, the access device can either be single-homed or dual-homed to aggregation devices.

Application of a Portstate Protection Group at Link Level

As shown in Figure 19-28, the active control board provides the working port and the standby control board provides the protection port. Such a Portstate protection group protects links.

Figure 19-28 Application of a link-level Portstate protection group



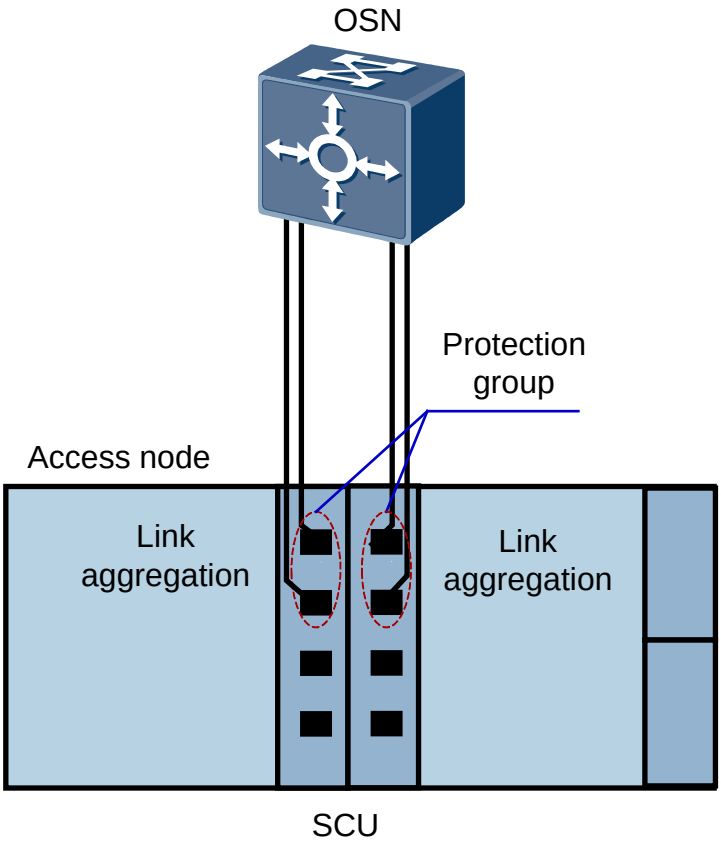
Application of a Portstate Protection Group at LAG Level

As shown in Figure 19-29, an LAG is configured on both the active and standby control boards. For the Portstate protection group, the LAG on the active control board is the working member, and the LAG on the standby control board is the protection member. Such a Portstate protection group protects LAGs and at the same time implements faster protection switching.

The following options are available for configuring LAGs on the MA5600T/MA5603T/MA5608T:

- You can configure multiple ports in one LAG. Select multiple ports on the active control board to form an LAG, and select multiple ports on the standby control board to form another LAG.
- You can also configure only one port in one LAG. Select one port on the active control board to form an LAG, and select one port on the standby control board to form another LAG.

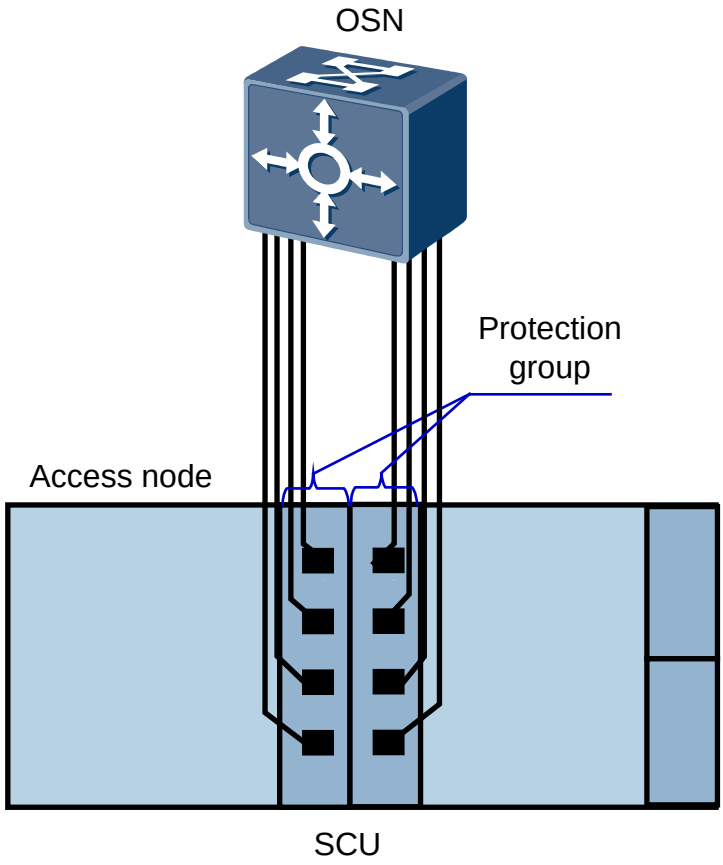
Figure 19-29 Application of an LAG-level Portstate protection group



Application of a Portstate Protection Group at Board Level

As shown in Figure 19-30, the active control board serves as the working member and the standby control board serves as the protection member. Such a Portstate protection group protects boards.

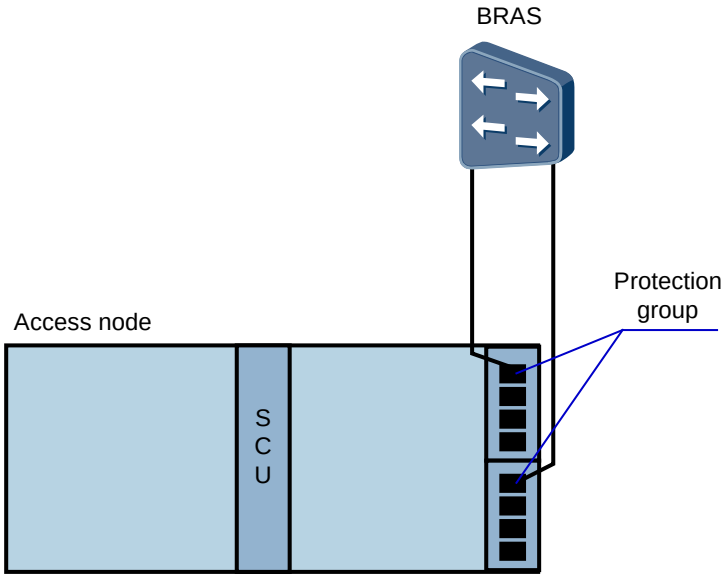
Figure 19-30 Application of a board-level Portstate protection group



Application of a Timedelay Protection Group (Upstream Single-Homing)

Figure 19-31 provides an example. In this example, two GIU boards each provide an upstream Ethernet port, which are connected to the same aggregation device. These two ports form a protection group to protect links.

Figure 19-31 Application of a Timedelay protection group (upstream single-homing)

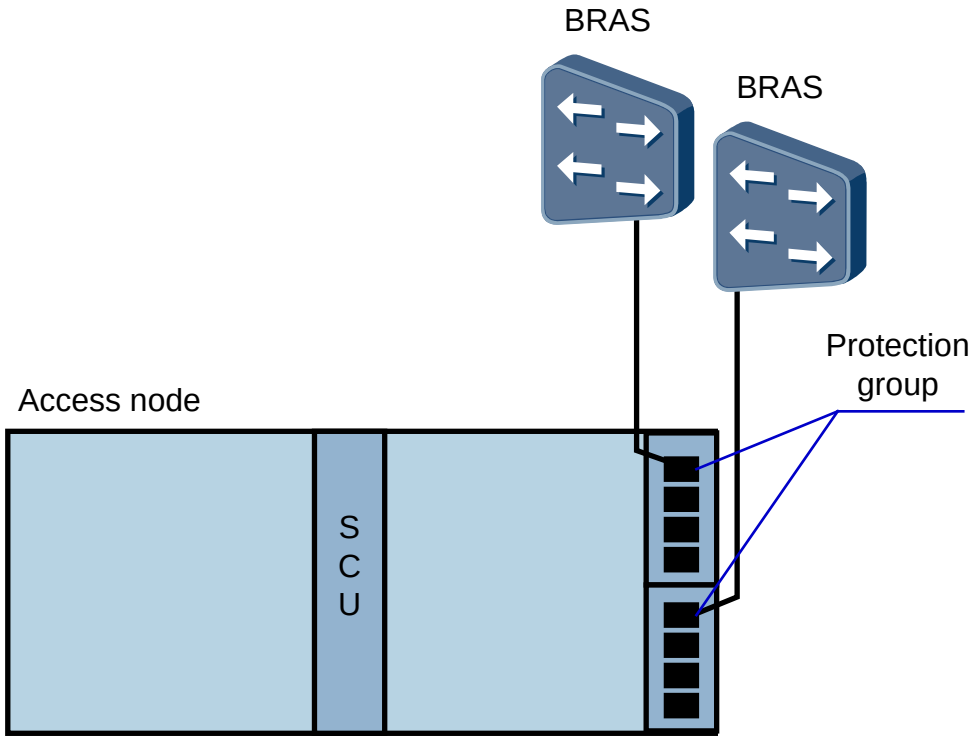


Application of a Timedelay Protection Group (Upstream Dual-Homing)

Figure 19-32 provides an example. In this example, two GIU boards each provide an upstream Ethernet port, which are connected to different aggregation devices. These two ports form a protection group to protect devices.

If the link between the access device and the master device (one of the dual-homed devices) fails, services are automatically switched to the protection link using the Layer 2 MAC address learning function. In this process, the two dual-homed devices do not need to perform switching.

Figure 19-32 Application of a Timedelay protection group (upstream dual-homing)



Typical Application of Protection Group and ARP Probe

If other devices, such as switch and transmission devices, are deployed between the access device and the aggregation devices, as shown in Figure 19-33, the status of links in the protection group cannot indicate whether the end-to-end links between the access device and the aggregation devices are normal. ARP probe helps resolve the issue by detecting the status of end-to-end links.

NOTE

The MA5600T/MA5603T/MA5608T does not support ARP probe for a protection group configured on the control board or an LAG-level protection group.

When ARP probe is used, ensure that devices over the ARP probe link between the source device and destination device do not terminate ARP probe packets.

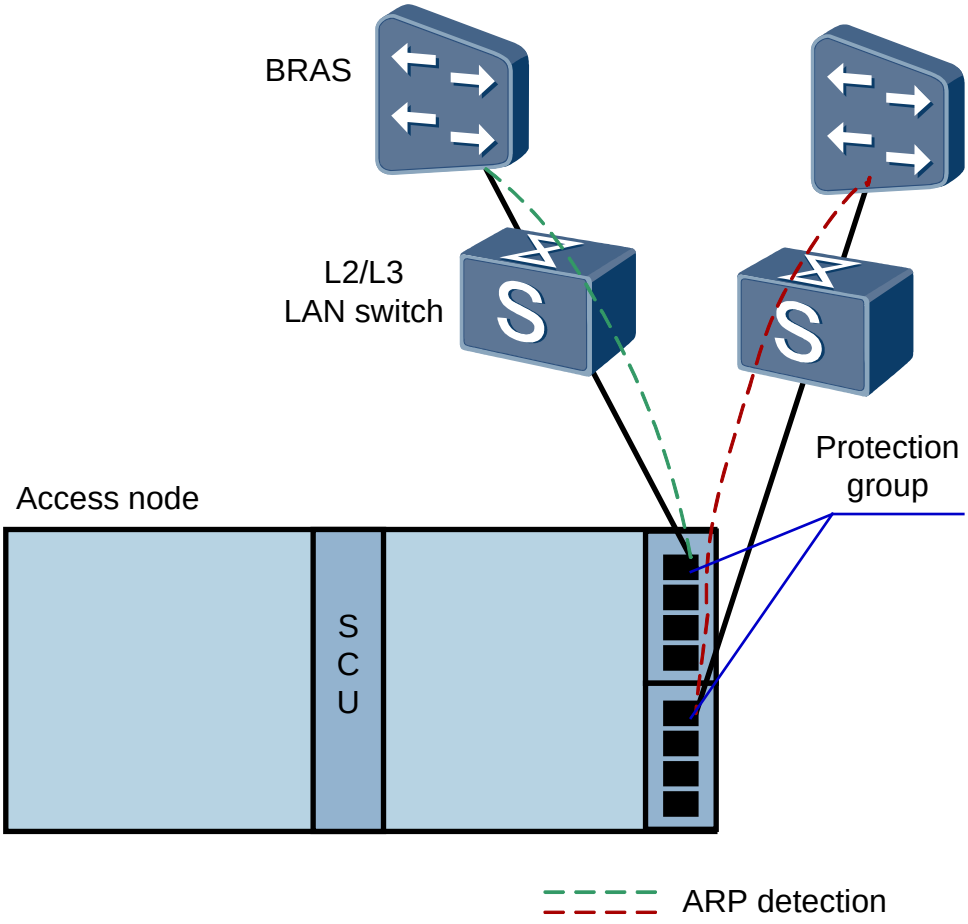
When ARP probe is used,

- The access device sends ARP request packets at specified intervals to request for the MAC address of the remote device, and considers the link normal if the remote device replies.
- If receiving no reply from the remote device within the specified waiting time, the access device considers the link faulty and triggers protection switching.

NOTE

The waiting time for ARP probe is equal to the interval for sending ARP request packets multiplied by the probe times. The minimum waiting time is 3s (1s x 3).

Figure 19-33 Typical application of protection group and ARP probe



19.4.5 Configuring an Ethernet Port Protection Group

The MA5600T/MA5603T/MA5608T allows two Ethernet ports to be bound together to provide protection. If the working port is faulty, the system switches services to the protection port. This ensures uninterrupted service forwarding and improves the reliability of links.

The MA5600T/MA5603T/MA5608T supports two types of Ethernet port protection groups:

- Portstate protection group: Applies to a network scenario in which the control board provides upstream ports, and protection is available at link level, link aggregation group (LAG) level, or board level. In the upstream direction, the access device is usually single-homed to a transmission device.
- Timedelay protection group: Applies to a network scenario in which the system control board or upstream board provides upstream ports, and protection is available only at link level. In the upstream direction, the access device can either be single-homed or dual-homed to aggregation devices.

Configuring a Link-Level Portstate Protection Group on the Control Board

A link-level Portstate protection group applies to the following scenario: The active and standby control boards on an access device each provide an upstream Ethernet port, and users want the upstream Ethernet port (working port) on the active control board to carry services

and the upstream Ethernet port (protection port) on the standby control board to back up. If the working port is faulty, an active/standby switchover is triggered, and the system switches services to the protection port to implement uninterrupted forwarding.

Prerequisites

- The boards and ports on the access device support Ethernet port protection groups. For details, see Feature Dependency and Limitation.
- In the protection group, a static MAC address is configured only on the working port. You can run the **display mac-address** command to query the static MAC address.
- The two ports on the interconnected device have the same data configurations and allow MAC address transfer (the same MAC address learned on different ports).

Procedure

Configure a link-level Portstate protection group on the control board.

Run the **protect-group** command to create a protection group (select **as-mainboard-port** as the protection level and **Portstate** as the protection type).

Step 1 Add a working port and a protection port to the protection group.

Run the **protect-group member** command to add a port on the control board as the working port and a port on the standby control board as the protection port in the protection group.

Step 2 Enable the protection group.

Run the **protect-group enable** command to enable the protection group.

Step 3 Query configurations about the protection group.

Run the **display protect-group** command to query configurations about the protection group. The configurations include the protection level, member ports, protection type, and reversion mode.

----End

Example

This example assumes a scenario in which two control boards on the MA5600T/MA5603T/MA5608T each provide an upstream port: 0/9/0 on the active control board and 0/10/0 on the standby control board. The two upstream ports form a link-level Portstate protection group.

NOTE
For details on the application diagram, see 19.4.4 Protection Group of Ethernet Ports Network Application.

To configure a link-level Portstate protection group in such a network scenario, do as follows:

NOTE

```
huawei(config)#protect-group 0 protect-target as-mainboard-port workmode portstate
huawei(protect-group-0)#protect-group member port 0/9/0 role work
huawei(protect-group-0)#protect-group member port 0/10/0 role protect
huawei(protect-group-0)#protect-group enable
huawei(protect-group-0)#display protect-group 0
  ------------------------------------------------------------------------
```

```
Group ID       : 0
Protect Target  : Port of active main board and standby main board
Work Mode      : portstate
Description    :
Admin State    : enable
Operation      : none
Reversion      : disable
Reversion Time(s): 720
--------------------------------------------------------------------------
Member       Role       Operation       State       PeerMember
--------------------------------------------------------------------------
0/9/0        work       none            active      none
0/10/0       protect    none            standby     none
--------------------------------------------------------------------------
```

Configuring an LAG-Level Portstate Protection Group on the Control Board

A link aggregation group (LAG)-level Portstate protection group applies to the following network scenario: The active and standby control boards on an access device each provide multiple upstream Ethernet ports, and users want the upstream Ethernet ports (working ports) on the active control board to carry services and those (protection ports) on the standby control board to back up. If a working port is faulty, and the LAG on the active control board has fewer properly-functioning ports than the LAG on the standby control board, an active/standby switchover is triggered, and the system switches services to the protection ports to implement uninterrupted forwarding.

Prerequisites

- The boards and ports on the access device support Ethernet port protection groups.
- In the protection group, a static MAC address is configured only on the working port. You can run the **display mac-address** command to query the static MAC address.
- The LAG ports on the interconnected device have the same data configurations and allow MAC address transfer.

Context

The combination of a Portstate protection group and an 19.3 Ethernet Link Aggregation ensures faster protection switching. The following options are available for configuring LAGs on the MA5600T/MA5603T/MA5608T:

- You can configure multiple ports in one LAG. Select multiple ports on the active control board to form one LAG, and select multiple ports on the standby control board to form another LAG.
- You can also configure only one port in one LAG. Select one port on the active control board to form one LAG, and select one port on the standby control board to form another LAG.

Note the following restrictions when you configure an LAG-level Portstate protection group on the control board:

- You need to configure an LAG first, and then add the master port of the LAG to the protection group.
- A port that is included in a protection group cannot be added to an LAG.

Procedure

Create an LAG on the active control board.

Run the **link-aggregation** command to create an LAG on the active control board. The LAG on the standby control board will be automatically created.

Step 1 Configure an LAG-level Portstate protection group on the control board.

Run the **protect-group** command to create a protection group (select **as-mainboard-lag** as the protection level and **Portstate** as the protection type).

Step 2 Add a working port and a protection port to the protection group.

Run the **protect-group member** command to add the master port in the LAG on the control board as the working port of the protection group, and the master port in the LAG on the standby control board as the protection port.

Step 3 Enable the protection group.

Run the **protect-group enable** command to enable the protection group.

Step 4 Query configurations about the protection group.

Run the **display protect-group** command to query configurations about the protection group. The configurations include the protection level, member ports, protection type, and reversion mode.

----End

Example

This example assumes a scenario in which two control boards on the MA5600T/MA5603T/MA5608T each provide two upstream ports to form LAGs: 0/9/0 and 0/9/1 on the active control board form one LAG, and 0/10/0 and 0/10/1 on the standby control board form the other LAG. The four upstream ports form one LAG-level Portstate protection group.

NOTE
For details on the application diagram, see 19.4.4 Protection Group of Ethernet Ports Network Application.

To configure an LAG-level Portstate protection group in such a network scenario, do as follows:

```
huawei(config)#link-aggregation 0/9 0-1 egress-ingress
huawei(config)#link-aggregation 0/10 0-1 egress-ingress
huawei(config)#protect-group 0 protect-target as-mainboard-lag workmode portstate
huawei(protect-group-0)#protect-group member port 0/9/0 role work
huawei(protect-group-0)#protect-group member port 0/10/0 role protect
huawei(protect-group-0)#protect-group enable
huawei(protect-group-0)#display protect-group 0
  ------------------------------------------------------------------------
  Group ID       : 0
  Protect Target  : LAG of active main board and standby main board
  Work Mode      : portstate
  Description    :
  Admin State    : enable
  Operation      : none
```

```
Reversion      : disable
Reversion Time(s): 720
--------------------------------------------------------------------------
Member       Role        Operation        State        PeerMember
--------------------------------------------------------------------------
0/9/0        work        none            active       none
0/10/0       protect      none            standby      none
--------------------------------------------------------------------------
```

Configuring a Board-Level Portstate Protection Group on the Control Board

A board-level Portstate protection group applies to the following network scenario: The active and standby control boards on the access device provide upstream ports, and users want the active control board (working board) to carry services and the standby one (protection board) to back up. If the working board is faulty, and the active control board has fewer properly-functioning ports than the standby control board, an active/standby switchover is triggered, and the system switches services to the protection board to implement uninterrupted forwarding.

Prerequisites

- The boards and ports on the access device support Ethernet port protection groups. For details, see Feature Dependency and Limitation.
- In the protection group, a static MAC address is configured only on the working port. You can run the **display mac-address** command to query the static MAC address.
- The ports on the interconnected device have the same data configurations and allow MAC address transfer.

Procedure

Configure a board-level Portstate protection group on the control board.

Run the **protect-group** command to create a protection group (select **as-mainboard** as the protection level and **Portstate** as the protection type).

Step 1 Add the working and protection boards to the protection group.

Run the **protect-group member** command to add the active control board as the working board of the protection group, and the standby control board as the protection board.

Step 2 Enable the protection group.

Run the **protect-group enable** command to enable the protection group.

Step 3 Query configurations about the protection group.

Run the **display protect-group** command to query configurations about the protection group. The configurations include the protection level, member ports, protection type, and reversion mode.

----End

Example

This example assumes a scenario in which the MA5600T/MA5603T/MA5608T connects to the upstream network through two control boards. The active control board 0/9 and the standby control board form a board-level Portstate protection group.

NOTE
For details on the application diagram, see 19.4.4 Protection Group of Ethernet Ports Network Application.

To configure a board-level Portstate protection group in such a network scenario, do as follows:

```
huawei(config)#protect-group 0 protect-target as-mainboard workmode portstate
huawei(protect-group-0)#protect-group member board 0/9 role work
huawei(protect-group-0)#protect-group member board 0/10 role protect
huawei(protect-group-0)#protect-group enable
huawei(protect-group-0)#display protect-group 0
  ------------------------------------------------------------------------
  Group ID       : 0
  Protect Target  : Active main board and standby main board
  Work Mode      : portstate
  Description    :
  Admin State    : enable
  Operation      : none
  Reversion      : disable
  Reversion Time(s): 720
  ------------------------------------------------------------------------
  Member      Role       Operation       State       PeerMember
  ------------------------------------------------------------------------
  0/9         work       none            active      none
  0/10        protect    none            standby     none
  ------------------------------------------------------------------------
```

Configuring a Timedelay Protection Group

A Timedelay protection group applies to the following scenario: The active and standby control boards or upstream service boards on an access device each provide an upstream Ethernet port, and users want port to carry services and the other port to back up. If the working port is faulty, the system switches services to the protection port to implement uninterrupted forwarding.

Prerequisites

- The boards and ports on the access device support Ethernet port protection groups. For details, see Feature Dependency and Limitation.
- In the protection group, a static MAC address is configured only on the working port. You can run the **display mac-address** command to query the static MAC address.
- In a protection group, the two interconnected ports have the same data configurations and allow MAC address transfer.

Procedure

(Optional) Configure the optical port shutdown function.

Run the **offline-tx-off-time** command to specify the time for keeping an optical port shut down in the case of a Linkdown. The optical port shutdown function helps improve protection switching performance.

Step 1 Create a Timedelay protection group.

Run the **protect-group** command to create a protection group (select **eth-nni-port** as the protection level and **Timedelay** as the protection type).

Step 2 Add a working port and a protection port to the protection group.

Run the **protect-group member** command to add one port on the control board or on the upstream board as the working port, and the other port on the board as the protection port in the protection group.

Step 3 Enable the protection group.

Run the **protect-group enable** command to enable the protection group.

Step 4 Query configurations about the protection group.

Run the **display protect-group** command to query configurations about the protection group. The configurations include the protection level, member ports, protection type, and reversion mode.

----End

Example

This example assumes a scenario in which the MA5600T/MA5603T/MA5608T connects to the upstream network through the GIU upstream service board. Two upstream ports 0/19/0 and 0/20/0 on the GIU board form a Timedelay protection group.

NOTE
For details on the application diagram, see 19.4.4 Protection Group of Ethernet Ports Network Application.

To configure a Timedelay protection group in such a network scenario, do as follows:

```
huawei(config)#interface giu 0/19
huawei(config-if-giu-0/19)#offline-tx-off-time 0 500
huawei(config)#quit
huawei(config)#interface giu 0/20
huawei(config-if-giu-0/20)#offline-tx-off-time 0 500
huawei(config)#quit
huawei(config)#protect-group 0 protect-target eth-nni-port workmode timedelay
huawei(protect-group-0)#protect-group member port 0/19/0 role work
huawei(protect-group-0)#protect-group member port 0/20/0 role protect
huawei(protect-group-0)#protect-group enable
huawei(protect-group-0)#display protect-group 0
  ------------------------------------------------------------------------
  Group ID       : 0
  Protect Target  : Port of Ethernet nni
  Work Mode      : timedelay
  Description    :
  Admin State    : enable
  Operation      : none
  Reversion      : disable
  Reversion Time(s): 720
```

```
-------------------------------------------------------------------------
Member       Role        Operation        State        PeerMember
-------------------------------------------------------------------------
0/19/0       work        none            active       none
0/20/0       protect      none            standby      none
-------------------------------------------------------------------------
```

19.5 Smart Link and Monitor Link

The smart link is a solution that is applied in the dual-upstream-transmission network and provides reliable and high-efficiency backup and quick switching for the dual uplinks. The monitor link solution, as a supplementary to the smart link solution, is used to monitor the uplinks.

19.5.1 Introduction to Smart Link and Monitor Link

Definition

The smart link is a solution that is applied in the network with dual uplinks and provides reliable and high-efficiency backup and quick switching for the dual uplinks.

Purpose

The network with dual uplinks is a common network application currently. In a network with dual uplinks, the redundant link can be blocked through the Spanning Tree Protocol (STP) or Rapid Spanning Tree Protocol (RSTP) and can provide the backup function. In this way, when the active link fails, traffic will be switched to the standby link.

The preceding two solutions (STP and RSTP) can meet customers' requirements for redundancy backup from the perspective of the function, but cannot meet the requirements of many users for the performance.

Thus, the smart link solution is applied to the access network. With this solution, redundancy backup for active and standby links and quick switching are implemented for a dual-homing network. This ensures high reliability and quick convergence. Meanwhile, as a supplementary to the smart link solution, the monitor link solution is introduced to monitor uplinks. This improves the backup function of the smart link solution.

Benefits

Benefits to Operators

Implementation of the smart link solution and the monitor link solution provides high reliability for carriers' network.

19.5.2 Smart Link

This topic describes the working principle of the smart link feature.

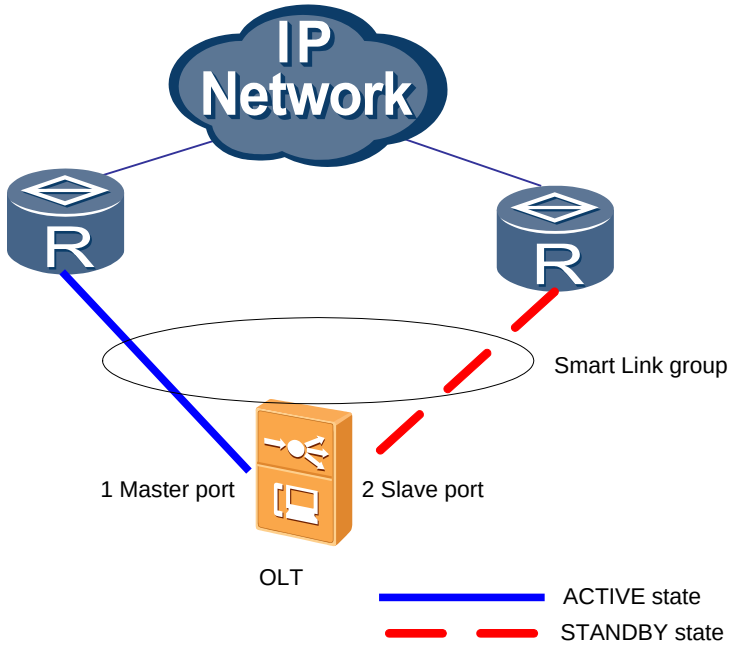
Basic Concepts

A smart link protect group can work in the following two modes:

- Active-standby working mode
- Load sharing working mode

Figure 19-34 shows the active-standby working mode of a smart link protect group.

Figure 19-34 Working mode of a smart link protect group



The following provides some concepts related to the smart link feature:

- Smart link group

  A smart link group is also called an intelligent link group, which contains up to two ports, namely, one master port and one slave port. In normal conditions, only one port is in the ACTIVE state, and the other port is blocked and in the STANDBY state. When the port in the ACTIVE state fails, the smart link group automatically blocks the port, and switches the previously standby port to the ACTIVE state. As shown in Figure 19-34, ports 1 and 2 form a smart link group.

- Master port

  The master port, which is also called the work port, is a port role in the smart link group. When both ports are in the STANDBY state, the master port is prevailed upon to switch to the ACTIVE state. The master port, however, is not always in the ACTIVE state. If the slave port is already in the ACTIVE state after link switching, the master port can only be in the STANDBY state even if its link recovers and the master port remains in this state until link switching the next time. For example, port 1 in the ACTIVE state in Figure 19-34 is the master port.

- Slave port

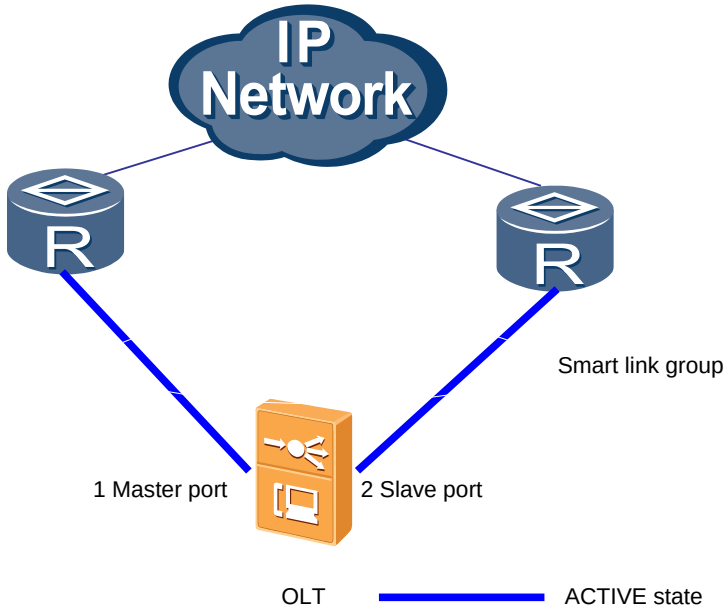
  The slave port, which is also called the protect port, is a port role in the smart link group. When both ports are in the STANDBY state, the master is prevailed upon to switch to the ACTIVE state, and the slave port remains in the STANDBY state. The slave port is not always in the STANDBY state. It switches to the ACTIVE state after link switching occurs on the master port. Port 2 in Figure 19-34 is the slave port.

- FLUSH packet

After link switching occurs on the smart link group, the original forwarding entry is not applicable to the network with new topology, and the upstream convergence device needs to update the MAC and ARP entries. In this case, the smart link group notifies the other devices on the network of updating the address table through sending the notification packet. This notification packet is the FLUSH packet.

Figure 19-35 shows the load sharing working mode of a smart link protect group.

Figure 19-35 Load sharing working mode of a smart link protect group

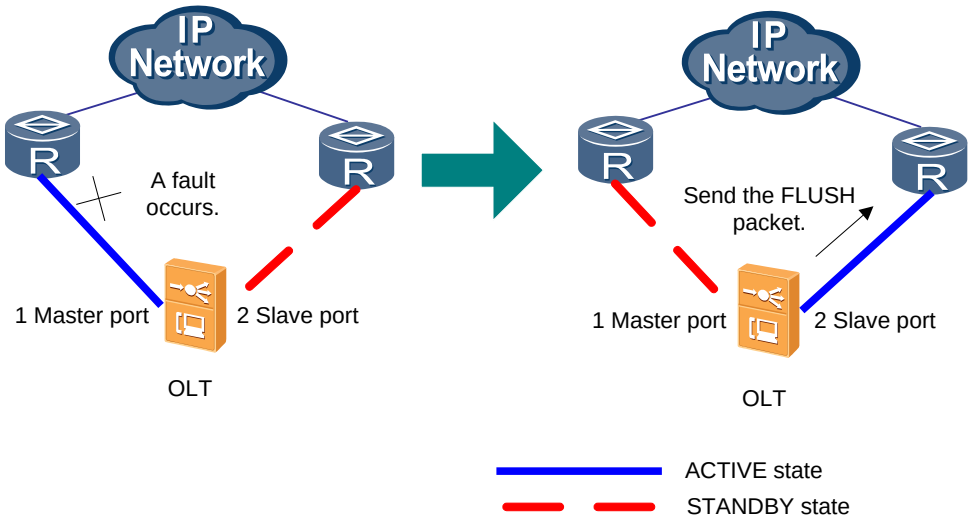


In the load sharing working mode, the links of both ports are enabled. If both ports are normal, some services are transmitted through the master port and the others are transmitted through the slave port. When either of the ports fails, all the services are transmitted through the port in the normal state.

Working Principle

Figure 19-36 shows the working principle of the smart link feature.

Figure 19-36 Working principle of the smart link feature



- Normal working state

  The link of port 1 on the device is the active link and the link of port 2 is the standby link. In normal conditions, port 1 is in the ACTIVE state and port 2 is in the STANDBY state.

- Switching

  When the link of port 1 fails, port 1 switches to the STANDBY state and port 2 switches to the ACTIVE state. When the original active link recovers from the fault, it remains in the blocked state and does not occupy bandwidth. This ensures stability of traffic.

- Update

  When link switching occurs in the smart link group, the MAC and ARP entries on the devices on the network may be incorrect. Therefore, a new mechanism for updating the MAC and ARP entries is required. Currently, there are the following two mechanisms available for updating the MAC and ARP entries:

  – The smart link device automatically updates the MAC and ARP entries through traffic.

  – The smart link device sends the FLUSH packet through the new link to update the MAC and ARP entries.

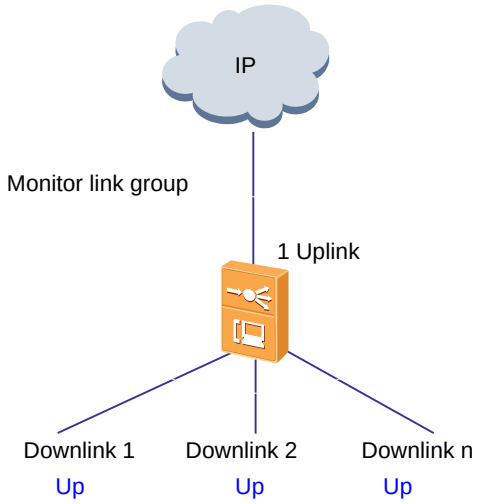
  When the device supports the first mechanism, bidirectional traffic trigger is required. This is applicable to the scenario when the device interoperates with the device from other vendors. When the device supports the second mechanism, it requires the upstream device to identify the FLUSH packet of smart link and to update the MAC and ARP entries.

19.5.3 Monitor Link

This topic describes the working principle of the monitor link feature.

Basic Concepts

Figure 19-37 Composition of a monitor link group



The following describes some basic concepts related to the monitor link feature.

- Monitor link group

  A monitor link group is composed of one uplink and several downlinks.

NOTE

The link in a monitor link group may not be a single link, but may be a certain type of link group. The uplink can be an aggregation group or protect group. The downlink can only be a single link. The status of the downlink changes according to the status of the uplink.

- Uplink

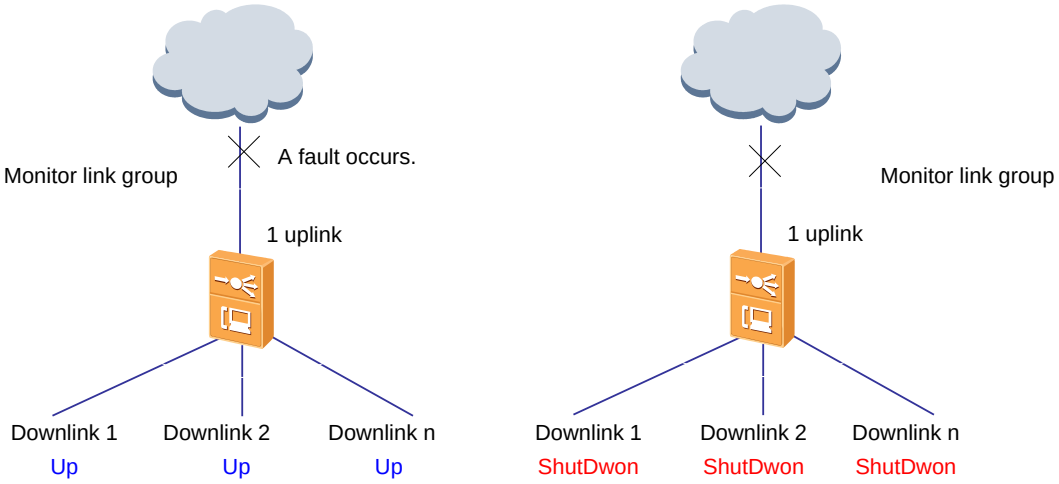
  When the uplink in a monitor link group fails, it indicates that the monitor link group fails. In this case, the downlinks in the monitor link group will be blocked by force.

- Downlink

  When a downlink in a monitor link group fails, it does not affect the uplink or the other downlinks.

Working Principle

Figure 19-38 Working principle of the monitor link feature

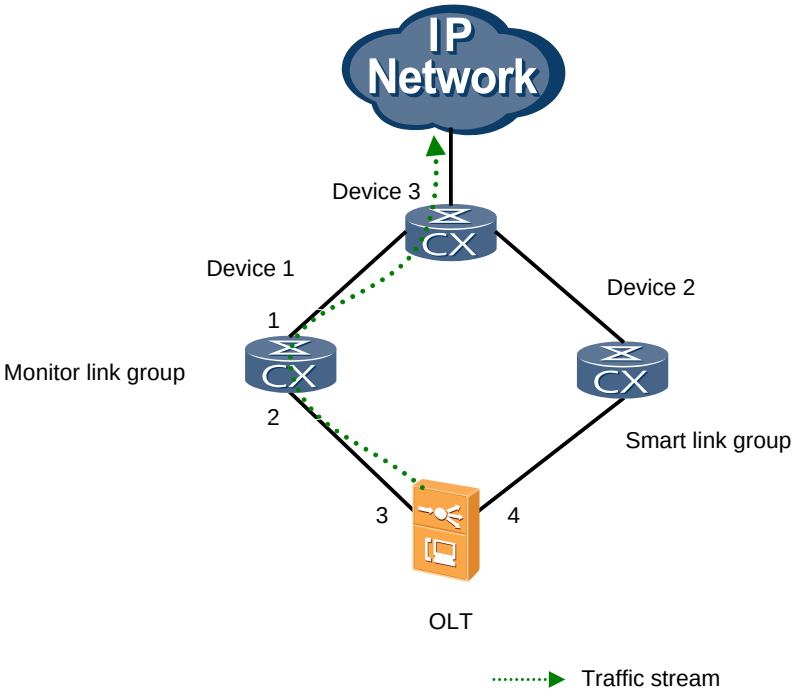


After a monitor link group is configured, its uplink will be monitored in real time. Once the uplink fails, all the UP downlinks in the monitor link group will be blocked by force. When the uplink recovers from the fault, the downlinks are resumed.

When the uplink is an aggregation group or protect group, the uplink is considered failed only when the entire aggregation group or protect group fails.

19.5.4 Smart Link and Monitor Link Network Applications

Figure 19-39 Network application of the smart link and monitor link features



The MA5600T/MA5603T/MA5608T works as the OLT. Ports 3 and 4 on the MA5600T/MA5603T/MA5608T are added to a smart link group and work in the active-standby mode. Port 1 on device 1 is configured as the upstream port of the monitor link group, and port 2 on device 1 the downstream port.

In normal conditions, traffic is transmitted through the path highlighted in green. If the uplink of device 1 fails, the uplink in the smart link group will be blocked. In this case, on the MA5600T/MA5603T/MA5608T, port 4 switches to the ACTIVE state because port 3 fails, and traffic is transmitted to device 2 and then to the upstream network.

If the monitor link group is not configured on device 1, the channel between device 1 and the MA5600T/MA5603T/MA5608T is still in the ACTIVE state when the channel between device 1 and device 3 fails. Thus, the user traffic will be transmitted to device 1 from the MA5600T/MA5603T/MA5608T. As a result, the user cannot access the network.

19.5.5 Configuring the Smart Link Redundancy Backup

The smart link is a solution that is applied in the network with dual uplinks and provides reliable and efficient backup and quick switching for the dual uplinks. The solution provides high reliability for carriers' network.

Context

Therefore, the smart link solution is applied to the access network. With this solution, redundancy backup for active and standby links and quick switching are implemented for a dual homing network. This ensures high reliability and quick convergence. Meanwhile, as a

supplementary to the smart link solution, the monitor link solution is introduced to monitor uplinks. This improves the backup function of the smart link solution.

The smart link and monitor link feature, which is applied to the scenario of a network with dual uplinks (the network is connected to the upstream IP network through dual uplinks), is related to the OLT and the upstream network device. The upstream network device such as the router must support the smart link and monitor link feature.

NOTE

The smart link and monitor link feature is put forth by Huawei. Currently, only Huawei devices support this technology.

Smart link-related concepts:

- Smart link protection group

  A smart link group contains up to two ports, namely one master port and one slave port. In normal conditions, only one port is in the active state, and the other port is blocked and in the standby state. When the port in the active state fails, the smart link group automatically blocks the port, and switches the previously standby port to the active state.

- Master port

  The master port, which is also called the work port, is a port role in a smart link group. When both ports are in the standby state, the master port takes priority to switch to the active state.

- Slave port

  The slave port, which is also called the protection port, is a port role in the smart link group. When both ports are in the standby state, the master is prevailed upon to switch to the active state, and the slave port remains in the standby state.

- Flush packet

  After link switching occurs on the smart link group, the original forwarding entry is not applicable to the network with new topology, and the upstream convergence device needs to update the MAC and ARP entries. In this case, the smart link group notifies the other devices in the network of updating the address table through sending the notification packet. This notification packet is the flush packet.

Monitor link-related concepts:

- Monitor link group

  A monitor link group is composed of one uplink and several downlinks.

- Uplink

  When the uplink in a monitor link group fails, the monitor link group fails. In this case, the downlinks in the monitor link group will be blocked by force.

- Downlink

  When a downlink in a monitor link group fails, it does not affect the uplink or the other downlinks.

A smart link can work in either the active/standby mode or the load balancing mode. The differences are as follows:

- In the active/standby mode, both ports are enabled. Only the master port is in the active state and can forward data. The slave port is blocked and is in the standby state.
- In the load balancing mode, both ports are enabled. If both ports work in the normal state, the data is forwarded through both ports, implementing load balancing.

Procedure

Configure a smart link protection group.

1. Run the **protect-group** command to create a smart link protection group. The protection group works in either the active/standby mode or the load balancing mode.

NOTE

- When configuring a smart link protection group, set the protected object to **eth-nni-port**. Working modes of other types do not support the smart link feature.
- Keyword **smart-link**: Indicates the smart-link active and standby mode. In this mode, both members in the PG are enabled, but only the active member forwards data.
- Keyword **smart-link load-balance**: Indicates the smart-link load balancing mode. In this mode, both links are enabled to share load to improve the usage ratio of the line.

2. Run the **protect-group member** command to add members to a smart link protection group.

   When adding members to the protection group, add a working member, and then add a protection member.

3. Run the **protect-group enable** command to enable the smart link protection group.

   After a protection group is created, the protection group is in the disabled state by default. You should enable the protection group to make the configuration take effect.

4. Query the information about the protection group.

   Run the **display protect-group** command to query the information about the protection group and all the members in the protection group.

Step 1 Configure the flush packet sending mode.

After service switching occurs on a protection group, the original forwarding entry is not applicable to the new network, and the entire network needs to update the MAC and ARP entries. In this case, the protection group sends flush packets to other devices to notify them of updating the MAC and ARP entries.

1. Run the **flush send** command to configure the flush packet sending parameters of the protection group, including the control VLAN and the password.
   a. If the flush packet sending parameters are not configured, no flush packet is sent when switching occurs on the protection group.
   b. If the protection group is not in the control VLAN, no flush packet is sent.
   c. The peer device must support receiving flush packets, and the flush packet receiving function of the corresponding port must be enabled.
2. Run the **display flush receive** command to query the port that receives flush packets and the flush packet receiving parameters.

Step 2 (Optional) Run the **load-balance instance** command to configure the load balancing parameters of a protection group.

Load balancing parameters determine that the working member and protection member carry different STP instances. Because VLANs are mapped to STP instances, the load balancing parameters in practice determine through which port (working member or protection member) the packets with different VLAN tags are transmitted.

NOTE

Configure the load balancing parameters only when the specified smart link protection group works in the load balancing mode.

- This command is used to configure STP instances that are carried by the protection member. The instances that are unconfigured are carried by the working member.

- The load balancing parameters of a protection group are based on STP instances pre-configured. You can run the **instance vlan** command to map VLANs to STP instances.

Step 3 (Optional) Configure a monitor link group.

The monitor link group and the smart link protect group are generally used together for monitoring the uplink and completing the smart link redundancy.

NOTE

1. Generally, the monitor link group is configured on the upper-layer device (such as a router) that is interconnected with the OLT, subtended to the smart link protection group.
2. You need to configure the monitor link on the MA5600T/MA5603T/MA5608T for monitoring the uplink of the subtended OLT only when the MA5600T/MA5603T/MA5608T functions as an upper-layer device interconnecting with the OLT. Otherwise, the configuration is meaningless.

1. Run the **monitor-link group** command to create a monitor link group, and enter the monitor link group mode.

   A monitor link group consists of one upstream port and multiple downstream ports. When the upstream port is faulty, the downstream ports are disabled. Therefore, the downstream devices can detect the link fault and switch the services to a normal link.
2. Run the **member port** command to add members to a monitor link group.
   - The uplink of a monitor link group can be a common Ethernet port, the master port of a protection group, or the master port of an aggregation group.
   - The downlink of a monitor link group can be only a common Ethernet port.
3. Run the **display monitor-link group** command to query the information about the monitor link group.

----End

Example

Assume the following configurations: The MA5600T/MA5603T/MA5608T implements dual uplinks through the GIU board, upstream ports 0/19/0 and 0/19/1 on the GIU board are added as members of smart link protection group 2, port 0/19/0 functions as the working port, port 0/19/1 functions as the protection port, the working mode is the load balancing mode, where,

- The STP instance 1 (mapping to VLAN 100-110) is carried by the working member.
- The STP instance 2 (mapping to VLAN 120-130) is carried by the protection member.
- The control VLAN of flush packets is VLAN 10, and the password is **abc**.

To perform these configurations and enable the protection group function, do as follows:

```
huawei(config)#stp region-configuration
huawei(stp-region-configuration)#instance 1 vlan 100 to 110
huawei(stp-region-configuration)#instance 2 vlan 120 to 130
huawei(stp-region-configuration)#active region-configuration
  STP actives region configuration,it may take several minutes,are you sure to
active region configuration? [Y/N][N]y
huawei(stp-region-configuration)#quit
huawei(config)#protect-group 2 protect-target eth-nni-port workmode smart-link
load-balance
huawei(config-protect-group-2)#protect-group member port 0/19/0 role work
huawei(config-protect-group-2)#protect-group member port 0/19/1 role protect
huawei(config-protect-group-2)#load-balance instance 2
huawei(config-protect-group-2)#flush send control-vlan 10 password simple abc
```

```
huawei(config-protect-group-2)#protect-group enable
huawei(config-protect-group-2)#quit
```

19.6 MSTP

The Multiple Spanning Tree Protocol (MSTP) is compatible with STP and RSTP.

19.6.1 Introduction to MSTP

Definition

The Spanning Tree Protocol (STP) applies to a loop network to realize path redundancy through certain algorithms. STP also prunes a loop network into a loop-free tree network. This helps to avoid proliferation and infinite loop of packets in the loop network.

The Rapid Spanning Tree Protocol (RSTP) is an improvement on STP. The rapidness of RSTP relies on the greatly shortened delay for the designated port and the root port to turn into the forwarding state in a certain condition. For details, see "Principle of RSTP" in "19.6.2 MSTP Principle." This helps to shorten the time for stabilizing the network topology.

The Multiple Spanning Tree Protocol (MSTP) is compatible with STP and RSTP.

Purpose

Although STP can prune a loop network into a loop-free network, it fails to transit fast. Even a port in a point-to-point link or an edge port has to wait double Forward Delay time before it can turn into the forwarding state.

RSTP features fast convergence; however, like STP, RSTP still has the following defects:

- All the bridges in a local area network (LAN) share a same spanning tree, and fail to block redundant links by VLAN.
- The packets of all the VLANs are forwarded along the same spanning tree. Therefore, load sharing of data traffic cannot be implemented between VLANs.

MSTP can be a remedy to the defects of STP and RSTP. It not only realizes fast convergence, but also enables traffic of different VLANs to be forwarded along their respective paths. This helps to provide a better load sharing mechanism for redundant links.

MSTP sets VLAN mapping tables (relation tables between VLANs and spanning trees) to associate VLANs and spanning trees. MSTP divides a switching network into multiple regions. Each region contains multiple spanning trees, and each spanning tree is independent from others.

MSTP prunes a loop network to a loop-free tree network to avoid proliferation and infinite loop of packets in the loop network. It also provides multiple redundant paths for data forwarding to realize load sharing of VLAN data during forwarding.

19.6.2 MSTP Principle

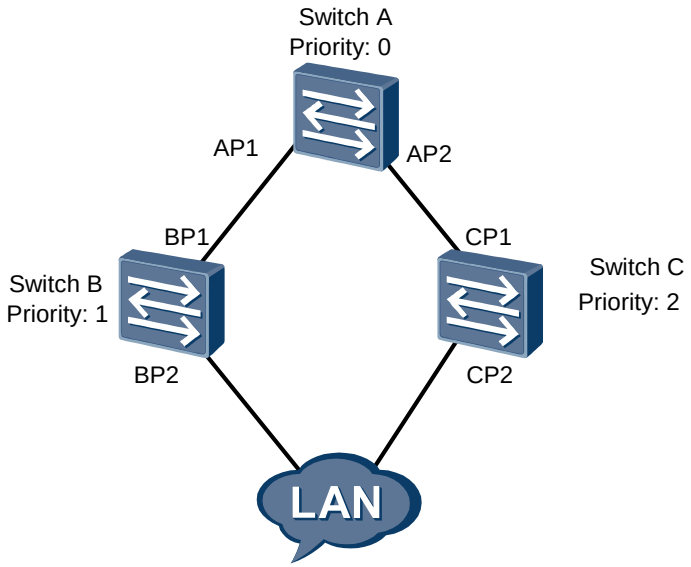
Principle of STP

STP determines the topology of a network by transmitting a certain special message (configuration message as defined in IEEE 802.1D) between bridges. A configuration message contains sufficient information to enable the bridge to complete the calculation of the spanning tree.

The following defines the designated port and the designated bridge:

- For a bridge (such as bridge A), the designated bridge is a bridge that is directly connected to bridge A and forwards data packets to bridge A. The designated port is the port in the designated bridge through which the data packets are forwarded to bridge A.
- For a LAN, the designated bridge is a bridge that forwards data packets to the LAN. The designated port is the port in the designated bridge through which the data packets are forwarded to the LAN.

Figure 19-40 Schematic drawing of designated bridge and designated port



As shown in Figure 19-40:

- AP1, AP2, BP1, BP2, CP1, and CP2 are ports in Switch A, Switch B, and Switch C respectively.
- Switch A forwards data to Switch B through port AP1, and then the designated bridge of Switch B is Switch A, and the designated port is port AP1 in Switch A.
- Switch B and Switch C are connected to the LAN. If Switch B forwards data packets to the LAN, the designated bridge of the LAN is Switch B, and the designated port is port BP2 in Switch B.

In STP, the configuration message is forwarded as follows:

1. In network initialization, all the bridges work as the root bridge of the spanning tree.
2. The designated port of a bridge takes the hello time as the interval for sending its configuration messages. If the port that receives the configuration message is a root port,

the bridge increases the message age contained in the configuration message by degrees and enables the timer to time the configuration message.

3. If a path fails, the root port on this path receives new configuration messages no longer, and the old configuration messages are discarded due to timeout. This results in recalculation of the spanning tree. A new path then is created to replace the faulty path and recover the network connectivity.

The new configuration message upon the recalculation, however, will not immediately spread throughout the entire network. In this case, the old root port and designated port that fail to discover the topology change will forward their data along the old paths. If the selected root port and designated port forwards data immediately, a temporary loop may be created.

Therefore, STP adopts a state transition mechanism. That is, the root port and the designated port have to experience a transition state before they can re-forward data. The transition state turns into the forwarding state upon Forward Delay. This delay guarantees that the new configuration message has spread throughout the entire network.

Defects of STP

- In case of topology change or link failure, a port has to wait double Forward Delay time before it can turn from the blocking state to the forwarding state. Therefore, in case of topology change, double Forward Delay time (at least scores of seconds) is required to restore the network connectivity.
- The entire bridged LAN uses a single spanning tree instance. Therefore, when the network is large, a longer convergence time may be required or the topology changes frequently.

Principle of RSTP

RSTP is an improvement on STP. The rapidness of RSTP relies on the greatly shortened delay for the designated port and the root port to turn into the forwarding state in a certain condition. This helps to shorten the time for stabilizing the network topology.

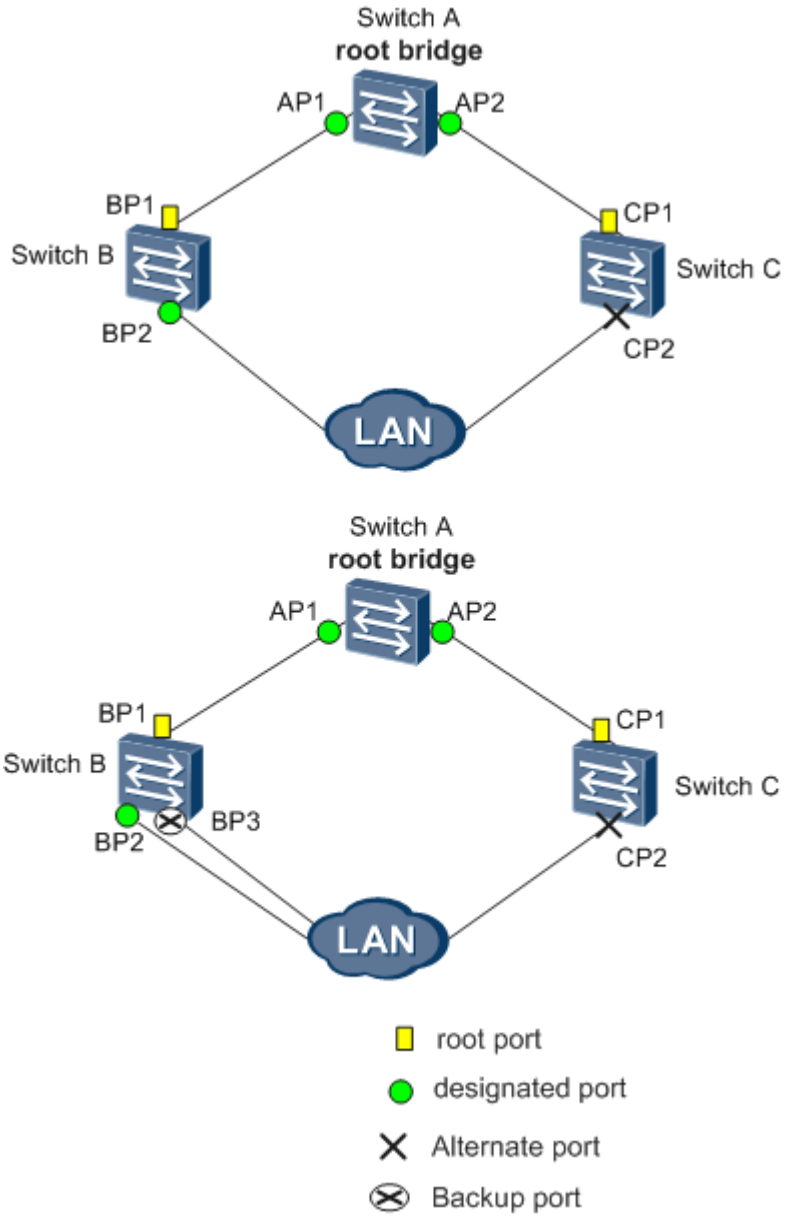
In comparison with STP, RSTP improves in the following aspects:

- First improvement:
  - The alternate port and backup port are set for rapid switching of the root port and designated port.
  - When the root port fails, the alternate port quickly switches to the new root port and turns into the forwarding state without delay.
  - When the designated port fails, the backup port quickly switches to the new designated port and turns into the forwarding state without delay.
- Second improvement:
  - In a point-to-point link connected with two switching ports, a designated port turns into the forwarding state without delay after one handshake with the downstream bridge.
  - In a shared link connected with at least three bridges, the downstream bridge does not respond to the handshake request sent from the upstream designated port, and the designated port has to wait double Forward Delay time before it turns into the forwarding state.
- Third improvement:

- A port that is directly connected to a terminal and is not connected to any other bridge is defined as an edge port. The edge port can directly turn into the forwarding state without delay.
- Because a bridge does not know whether a port is directly connected to a terminal, the edge port must be configured manually.

RSTP defines four port roles: root port, designated port, alternate port, and backup port. As shown in Figure 19-41.

Figure 19-41 Diagram of port roles



The functions of the root port and designated port are the same as those defined in STP. The alternate port and backup port are described as follows:

- From the perspective of configuration BPDU transmission:

  - An alternate port is blocked after learning the configuration BPDUs sent by other bridges.
  - A backup port is blocked after learning the configuration BPDUs sent by itself.
- From the perspective of user traffic
  - An alternate port backs up the root port and provides an alternate path from the designated bridge to the root bridge.
  - A backup port backs up the designated port and provides an alternate path from the root node to the leaf node.

  After all RSTP-capable ports are assigned roles, topology convergence is completed.

The bridges that adopt RSTP are compatible with the bridges which adopt STP. The bridges that adopt RSTP can identify both STP and RSTP packets and apply them to calculation of the spanning tree.

Defects of RSTP

Although RSTP features fast convergence, like STP, RSTP still has the following defects: All the bridges in a LAN share a same spanning tree, and thus the packets of all the VLANs cannot be forwarded equally. Furthermore, the packets of some VLANs cannot be forwarded.

Principle of MSTP

MSTP can compensate for the defects of STP and RSTP. It not only realizes fast convergence, but also enables traffic of different VLANs to be forwarded along their respective paths. This helps to provide a better load sharing mechanism for redundant links.

MSTP sets VLAN mapping tables (relation tables between VLANs and spanning trees) to associate VLANs and spanning trees. MSTP divides a switching network into multiple regions. Each region contains multiple spanning trees, and each spanning tree is independent of one another. Multiple spanning trees can run on each bridge to forward the packets of different VLANs.

MSTP divides the entire Layer 2 network into multiple spanning tree (MST) regions. These regions and the other bridges and LANs are connected into a single common spanning tree (CST). Multiple spanning trees are created in a region through calculation. Each spanning tree is defined as a multiple spanning tree instance (MSTI). MSTI 0 is defined as an internal spanning tree (IST). MSTP connects all bridges and LANs with a single common and internal spanning tree (CIST) which consists of the CST and the IST. Like RSTP, MSTP calculates the spanning tree according to the configuration message. The configuration message, however, contains the message of MSTP on the bridge.

- Calculation of CIST
  - Select a bridge with the highest priority within the entire network as the CIST root by comparing the configuration messages.
  - In each MST region, MSTP creates an IST through calculation. Meanwhile, MSTP regards each MST region as a single bridge, and then creates a CST between regions.
  - The CST and the IST forms the CIST that connects all the bridges in a bridge network.
- Calculation of MSTI

  In an MST region, MSTP creates different MSTIs for different VLANs according to the mapping relation between the VLANs and the spanning tree instances. Each spanning

tree is calculated independently. The process is similar to that in which the RSTP calculates the spanning tree.

Based on RSTP, MSTP has two additional port types. MSTP ports can be root ports, designated ports, alternate ports, backup ports, edge ports, master ports, and regional edge port.

The functions of root ports, designated ports, alternate ports, backup ports, and edge ports have been defined in RSTP. Table 19-8 lists all port roles in MSTP.

NOTE

Except edge ports, all ports participate in MSTP calculation.

A port can play different roles in different spanning tree instances.

Table 19-8 Port roles

Port Role	Description
Root port	A root port is the non-root bridge port closest to the root bridge. Root bridges do not have root ports. Root ports are responsible for sending data to root bridges. As shown in Figure 19-42, Switch A is the root; BP1 is the root port on Switch B; CP1 is the root port on Switch C.
Designated port	The designated port on a switching device forwards BPDUs to the downstream switching device. As shown in Figure 19-42, AP1 and AP2 are designated ports on Switch A; BP2 is a designated port on Switch B.
Alternate port	• From the perspective of sending BPDUs, an alternate port is blocked after a BPDU sent by another bridge is received. • From the perspective of user traffic, an alternate port provides an alternate path to the root bridge. This path is different than using the root port. As shown in Figure 19-42, CP2 is an alternate port.
Backup port	• From the perspective of sending BPDUs, a backup port is blocked after a BPDU sent by itself is received. • From the perspective of user traffic, a backup port provides a backup/redundant path to a segment where a designated port already connects. As shown in Figure 19-42, BP3 is a backup port.
Master port	A master port is on the shortest path connecting MST regions to the CIST root. BPDUs of an MST region are sent to the CIST root through the master port. Master ports are special regional edge ports, functioning as root ports on ISTs or CISTs and master ports in instances. As shown in Figure 19-43, Switch A, Switch B, Switch C, and Switch D form an MST region. AP1 on Switch A, being the nearest port in the region to the CIST root, is the master port.
Regional edge port	A regional edge port is located at the edge of an MST region and connects to another MST region or an SST. During MSTP calculation, the roles of a regional edge port in the MSTI and

Port Role	Description
	the CIST instance are the same. If the regional edge port is the master port in the CIST instance, it is the master port in all the MSTIs in the region. As shown in Figure 19-43, AP1, DP1, and DP2 in an MST region are directly connected to other regions, and therefore they are all regional edge ports of the MST region. AP1 is a master port in the CIST. Therefore, AP1 is the master port in every MSTI in the MST region.
Edge port	An edge port is located at the edge of an MST region and does not connect to any switching device. Generally, edge ports are directly connected to terminals.

Figure 19-42 Root port, designated port, alternate port, and backup port

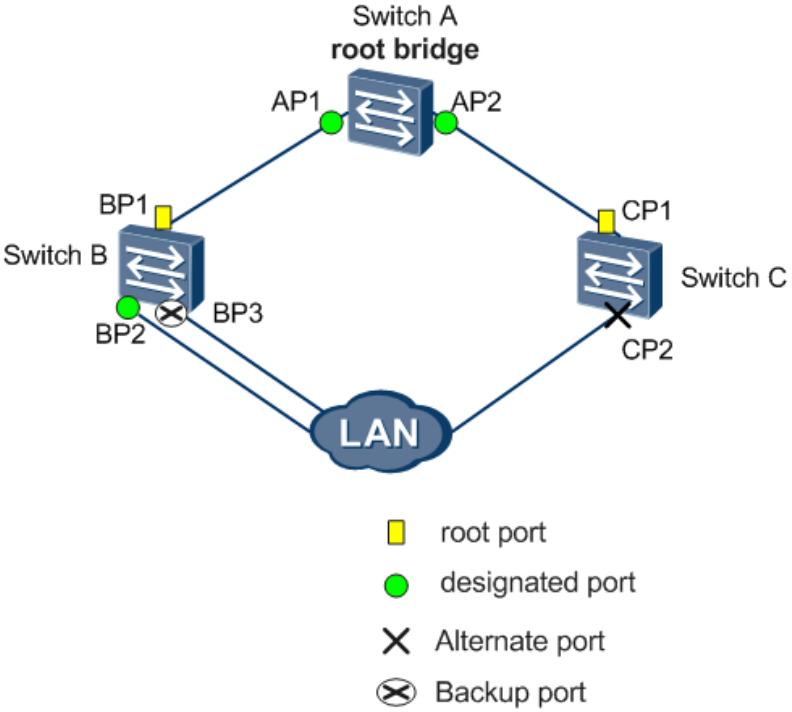
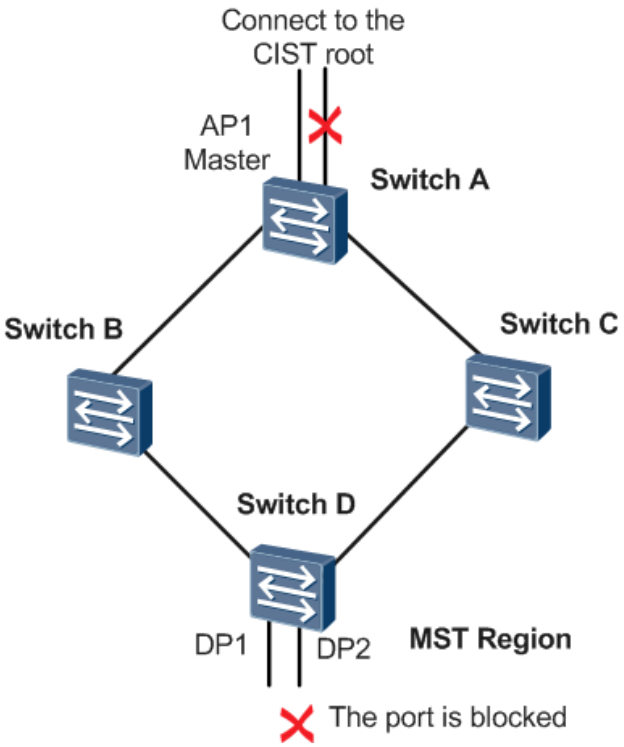


Figure 19-43 Master port and regional edge port



Implementation of MSTP on the MA5600T/MA5603T/MA5608T

MSTP is compatible with STP and RSTP. The bridges that adopt MSTP can identify both STP and RSTP packets and apply them to calculate spanning tree.

Besides the basic functions of MSTP, the MA5600T/MA5603T/MA5608T provides some special functions, such as:

- BPDU protection

  For an access device, the access port is generally connected to a terminal (such as a PC) or file server. In this case, the access port is set to an edge port for the purpose of fast transition. When receiving a configuration message (BPDU), the edge port switches to a non-edge port automatically, the spanning tree is re-calculated and the topology changes accordingly.

  In normal conditions, an edge port cannot receive STP configuration messages. If the bridge is maliciously attacked by forged configuration messages, the network will be attacked. The BPDU protection function can prevent such network attacks.

  After the BPDU protection function is enabled on the MA5600T/MA5603T/MA5608T, if an edge port receives a configuration message, the system shuts down the edge port, and notifies the network management system of the related information. Only network administrators can enable the port that is shut down.

  It is recommended that you enable the BPDU protection function on the MA5600T/MA5603T/MA5608T which is configured with an edge port.

- Root protection

  Because of wrong configurations by the maintenance personnel or malicious network attacks, a legal root bridge in the network may receive a configuration message with a

higher priority. In this case, this root bridge may become a non-root bridge and the topology changes accordingly. Such illegal change results in transfer of traffic in high-speed links to low-speed links, thus causing network congestion.

The root protection function is a solution to this problem.

When the root protection function is enabled for a port, the port is always a designated port. If the port receives a configuration message with a higher priority, and is to become a non-designated port, the port will turn into the listening state and will not forward packets (that is, the link connected to the port is disconnected). If the port does not receive a configuration message of a much higher priority within a certain long period of time, the port will turn into the normal state.

- Loop protection

  A bridge maintains the states of the root port and other blocked ports by continuously receiving BPDUs from the upstream bridge.

  In case of link congestion or failure, these ports fail to receive BPDUs from the upstream bridge. For this reason, the bridge will re-select its root bridge. The previous root bridge switches to the designated port, and the blocked ports turn to the forwarding state. As a result, loops are created in the switching network.

  The loop protection function is a solution to this problem.

  After receiving the BPDUs (excluding the TCN packets) again, a port under loop protection normally processes the packets, selects the role, and resets the forwarding state of the port. The port is not always in the blocked state.

  When the loop protection function is enabled, if the root port switches to a non-root port, it will turn into the discarding state, and the blocked ports will remain in the discarding state. Therefore, no packets are forwarded, and no loop is created in the network.

NOTE

The three protection functions conflict with each other.

19.6.3 Configuring the MSTP

The MA5600T/MA5603T/MA5608T supports the application of the Multiple Spanning Tree Protocol (MSTP), Spanning Tree Protocol (STP), and Rapid Spanning Tree Protocol (RSTP). The MA5600T/MA5603T/MA5608T supports the MSTP ring network, which can meet various networking requirements.

Context

- MSTP applies to a redundant network. It makes up for the drawback of STP and RSTP. MSTP makes the network converge fast and the traffic of different VLANs distributed along their respective paths, which provides a better load-sharing mechanism.
- MSTP trims a loop network into a loop-free tree network. It prevents the proliferation and infinite cycling of the packets in the loop network. In addition, MSTP supports load sharing by VLAN during data transmission.
- The status of the transparent transmission for bridge protocol data unit (BPDU) packets configured by running the **bpdu tunnel** command determines the destination MAC address of MSTP packets.
  - If the transparent transmission of BPDU packets is disabled in all VLANs, the destination MAC address of MSTP packet is 01:80:C2:00:00:00.
  - If the transparent transmission of BPDU packets is enabled in any VLAN, the destination MAC address of MSTP packet is 01:80:C2:00:00:08.

Procedure

Enabling the MSTP function.

- By default, the MSTP function is disabled.
- After the MSTP function is enabled, the device determines whether it works in STP compatible mode or MSTP mode based on the configured protocol.
- After the MSTP function is enabled, MSTP maintains dynamically the spanning tree of the VLAN based on the received BPDU packets. After the MSTP function is disabled, the MA5600T/MA5603T/MA5608T becomes a transparent bridge and does not maintain the spanning tree.

1. Run the **stp enable** command to enable the MSTP function of the bridge.
2. Run the **stp port enable** command to enable the MSTP function of the port.
3. Run the **display stp** command or the **display stp port** command to query the MSTP state of the bridge or the port.

Step 1 Configuring the MST region name.

1. Run the **stp region-configuration** command to enter MST region mode.
2. Run the **region-name** command to configure the name of the MST region.

   By default, the MST region name is the bridge MAC address of the device.

Step 2 Configuring the MSTP instance.

The MSTP protocol configures the VLAN mapping table (mapping between the VLAN and the spanning tree), which maps the VLAN to the spanning tree.

1. Run the **stp region-configuration** command to switch over to MST region mode.
2. Run the **instance vlan** command to map the specified VLAN to the specified MSTP instance.
   - By default, all VLANs are mapped to CIST, that is, instance 0.
   - One VLAN can be mapped to only one instance. If you re-map a VLAN to another instance, the original mapping is disabled.
   - A maximum of 10 VLAN sections can be configured for an MSTP instance.

   NOTE
   A VLAN section refers to the consecutive VLAN IDs from the start VLAN ID to the end VLAN ID.
3. Run the **check region-configuration** command to query the parameters of the current MST region.

Step 3 Activating the configuration of the MST region.

1. Run the **stp region-configuration** command to switch over to MST region mode.
2. Run the **active region-configuration** command to activate the configuration of the MST region.
3. Run the **display stp region-configuration** command to query the effective configuration of the MST region.

Step 4 Setting the priority of the device in the specified spanning tree instance.

1. Run the **stp priority** command to set the priority of the device in the specified spanning tree instance.
2. Run the **display stp** command to query the MSTP configuration of the device.

Step 5 Other optional configurations.

- Setting the MST region parameters.
  - Run the **stp md5-key** command to set the MD5-Key for the MD5 encryption algorithm configured on the MST region.
  - In the MSTP region mode, run the **vlan-mapping module** command to map all VLANs to the MSTP instances by modular arithmetic.
  - In the MSTP region mode, run the **revision-level** command to set the MSTP revision level of the device.
  - Run the **reset stp region-configuration** command to restore the default settings to all parameters of the MST region.
- Specifying the device as a root bridge or a backup root bridge.
  - Run the **stp root** command to specify the device as a root bridge or a backup root bridge.
- Setting the time parameters of the specified network bridge.
  - Run the **stp timer forward-delay** command to set the Forward Delay of the specified network bridge.
  - Run the **stp timer hello** command to set the Hello Time of the specified network bridge.
  - Run the **stp timer max-age** command to set the Max Age of the specified network bridge.
  - Run the **stp time-factor** command to set the timeout time factor of the specified network bridge.
- Setting the parameters of the specified port.
  - Run the **stp port transmit-limit** command to set the number of packets transmitted by the port within the Hello Time.
  - Run the **stp port edged-port enable** command to set the port as an edge port.
  - Run the **stp port cost** command to set the path cost of a specified port.
  - Run the **stp port port-priority** command to set the priority of the specified port.
  - Run the **stp port point-to-point** command to set whether the link that is connected to the port is a point-to-point link.
- Configuring the device protection function.
  - Run the **stp bpdu-protection enable** command to enable the BPDU protection function of the device.
  - Run the **stp port loop-protection enable** command to enable the loop protection function of the port.
  - Run the **stp port root-protection enable** command to enable the root protection function of the port.
- Setting the maximum number of hops of the MST region.
  - Run the **stp max-hops** command to set the maximum number of hops of the MST region.
- Setting the diameter of the switching fabric.
  - Run the **stp bridge-diameter** command to set the diameter of the switching fabric.
- Setting the calculation standard for the path cost.
  - Run the **stp pathcost-standard** command to set the calculation standard for the path cost.
- Clear the MSTP protocol statistics.
  - Run the **reset stp statistics** command to clear the MSTP protocol statistics.

----End

Example

Configure the MSTP parameters as follows:

- Enable the MSTP function.
- Enable the MSTP function on port 0/19/0.
- Set the MSTP running mode to MSTP compatible mode.
- Configure MST region parameters:
  - Configure the MD5-Key for the MD5 encryption algorithm to 0x11ed224466.
  - Configure the MST region name to huawei-mstp-bridge.
  - Map VLAN2-VLAN10 and VLAN12-VLAN16 to MSTP instance 3.
  - Map all the VLANs to the specified MSTP instances.
  - Configure the MSTP revision level of the device to 100.
- Configure the maximum hops for the MST region to 10.
- Activate the configuration of the MST region manually.
- Configure the priority of the device in spanning tree instance 2 to 4096.
- Configure the current device as the root bridge of MSTP instance 2.
- Configure the diameter of the switching network to 6.
- Configure the calculation standard for the path cost to IEEE 802.1t.
- Configure the time parameters of a specified bridge:
  - Configure the forward delay to 2000 centiseconds.
  - Configure the hello time to 1000 centiseconds.
  - Configure the max age to 3000 centiseconds.
  - Configure the timeout time factor to 6.
- Configure the parameters of a specified port:
  - Configure the maximum number of packets transmitted in a hello time period to 16.
  - Configure port 0/19/0 to be an edge port.
  - Configure the path cost of the port in a specified spanning tree instance to 1024.
  - Configure the priority of the port to 64.
  - The link connected to port 0/19/0 is a point-to-point link.
- Enable the BPDU protection function on the device.

```
huawei(config)#stp enable
  Change global stp state may active region configuration,it may take several
minutes,are you sure to change global stp state? [Y/N][N]y
huawei(config)#stp port 0/19/0 enable
huawei(config)#stp mode mstp
huawei(config)#stp md5-key 11ed224466
huawei(config)#stp region-configuration
huawei(stp-region-configuration)#region-name huawei-mstp-bridge
huawei(stp-region-configuration)#instance 3 vlan 2 to 10 12 to 16
huawei(stp-region-configuration)#vlan-mapping module 16
huawei(stp-region-configuration)#revision-level 100
huawei(stp-region-configuration)#active region-configuration
huawei(stp-region-configuration)#quit
```

```
huawei(config)#stp instance 2 priority 4096
huawei(config)#stp instance 2 root primary
huawei(config)#stp max-hops 10
huawei(config)#stp bridge-diameter 6
huawei(config)#stp pathcost-standard dot1t
huawei(config)#stp timer forward-delay 2000
huawei(config)#stp timer hello 1000
huawei(config)#stp timer max-age 3000
huawei(config)#stp time-factor 6
huawei(config)#stp port 0/19/0 transmit-limit 16
huawei(config)#stp port 0/19/0 edged-port enable
huawei(config)#stp port 0/19/0 instance 0 cost 1024
huawei(config)#stp port 0/19/0 instance 0 port-priority 64
huawei(config)#stp port 0/19/0 point-to-point force-true
huawei(config)#stp bpdu-protection enable
```

19.6.4 MSTP Reference Standards and Protocols

The following lists the reference documents of MSTP:

- IEEE Std 802.1d, 1998 Edition, Spanning Tree Protocol
- IEEE Std 802.1w-2001, Rapid Spanning Tree Protocol
- IEEE Std 802.1s-2002, Multiple Spanning Tree Protocol

19.7 RRPP

Rapid Ring Protection Protocol (RRPP) is a link-layer protocol specially used for protecting Ethernet ring networks.

19.7.1 Introduction to RRPP

Definition

Most metropolitan area networks (MANs) and enterprise networks adopt a ring topology to provide high reliability. In a ring topology, the failure of any node on the ring does not affect services. The following introduces some known ring network technologies.

- SDH/SONET ring

  Synchronous digital hierarchy (SDH) and synchronous optical network (SONET) are ring technologies widely used in current transport networks and support single ring and multiple rings. SDH/SONET feature high reliability because they provide an automatic protection switching (APS) self-healing mechanism in case of a fault.

  Due to the point-to-point (P2P) and circuit-switched design, in SDH/SONET ring networks, bandwidth is fixedly allocated and reserved on the P2P links between nodes. Thus bandwidth cannot be adjusted according to actual traffic condition in the networks. This hampers the efficient utilization of bandwidth and makes it different for the SDH/SONET networks to adapt to IP data service, which has the bursty characteristics.

  Broadcast and multicast packets in SDH/SONET ring networks are fragmented and transmitted as multiple unicast packets, which is a serious waste of bandwidth. In addition, a redundant bandwidth as high as 50% is required for the APS feature. In this case, a flexible selection mechanism is not available.

- RPR ring

  Resilient packet ring (RPR) is a MAC layer-based protocol researched and standardized by the IEEE 802.17 working group and RPR Alliance. RPR is used on ring topologies. The RPR design targets at a close-loop, P2P, and MAC layer-based logical ring topology.

  Viewed from the physical layer, an RPR is a set of P2P links; from the data link layer, RPR is more like a broadcast medium network similar to Ethernet.

  RPR requires dedicated hardware support and involves complicated fairness algorithms.

- STP ring

  Spanning Tree Protocol (STP) is also a standard ring protection protocol developed by IEEE and has been in wide application. However, STP rings in actual application are restricted by the network scale, and the convergence time is also subject to the network topology. The convergence time is not desirable when the network diameter is large. In this case, STP rings may fail to carry data that has high requirements on transmission quality.

Rapid Ring Protection Protocol (RRPP) is a link-layer protocol dedicated to Ethernet ring protection. RRPP is free from the problems above, such as bandwidth waste, dedicated hardware support, and slow convergence. On a complete Ethernet ring RRPP protects against broadcast storms caused by data loops. When the Ethernet ring has a link break, RRPP can rapidly recover the communication channels between the nodes on the ring.

Purpose

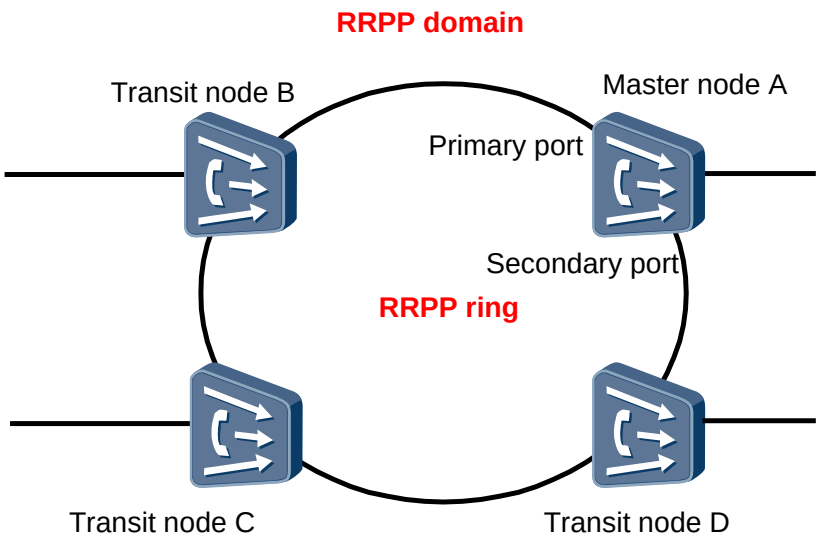
To enable faster convergence and mitigate the impact of network scale on the convergence speed, Huawei develops RRPP, a link-layer protocol specially for Ethernet ring protection. Compared with other Ethernet ring network technologies, RRPP has the following advantages:

- Provides fast topology convergence within 50 ms.
- Supports a convergence duration independent from the number of nodes on the ring. RRPP applies to networks with a larger diameter.
- Prevents broadcast storms caused by data loops when the Ethernet ring is complete.
- Rapidly starts the backup link to recover the communication channel between the nodes on the Ethernet ring when the ring has a link break.

19.7.2 RRPP Network Topology

An RRPP domain has the following constituents, as shown in Figure 19-44.

Figure 19-44 RRPP domain



RRPP Domain

An RRPP domain is uniquely identified by an integral ID, and it consists of a set of interoperated switches that are configured with the same domain ID and the same control VLANs. One node supports only one domain.

An RRPP domain mainly includes the following components:

- RRPP ring
- Control VLAN
- Master node
- Transit node

RRPP Ring

An RRPP ring physically corresponds to a ring-connection Ethernet topology. An RRPP domain is built on multiple interconnected RRPP rings, among which there is a primary ring and the rest are secondary rings. The primary ring and secondary rings are identified by levels specified during configuration. The primary ring is identified by level 0 and secondary rings by level 1.

An RRPP ring is also identified by an integral ID. Currently, the MA5600T supports only one RRPP ring in an RRPP domain.

Control VLAN

Each RRPP domain has two control VLANs. One is the primary control VLAN and the other the secondary control VLAN. The protocol packets of the primary ring are transmitted in the primary control VLAN, and the protocol packets of the secondary ring are transmitted in the secondary control VLAN.

During configuration, you only need to specify the ID of the primary control VLAN, and a VLAN whose ID is larger than the primary control VLAN by 1 will serve as the secondary control VLAN. The ports of the primary control VLAN and the secondary control VLAN must not be configured with IP addresses.

The RRPP port on the primary ring must belong to the primary control VLAN and the secondary control VLAN at the same time; the RRPP port of the secondary ring only needs to belong to the secondary control VLAN.

The primary ring is treated as a logical node of the secondary rings and the packets of the secondary rings are transparently transmitted by the primary ring. The packets of the primary ring are transmitted only within the primary ring and are not transmitted to the secondary rings.

Master Node

The master node is the policy-making and controlling node on an RRPP ring. Each RRPP ring must have one and only one designated master node. The master node initiates the polling mechanism (a mechanism for actively checking the ring status), and also determines and implements the policies after the network topology changes.

A master node has three states:

- Complete state: If the master node can receive its own hello packets on the secondary port, it indicates that the ring is complete. In this case, the ring is in the complete state.
- Failed state: If the master node does not receive its own hello packet within a specified time, the master node regards that there is a link-down on the ring network. In this case, the master node opens its secondary port for forwarding data and the ring is in the failed state.
- Unknown state: When the RRPP ring is not enabled, the ring is in the unknown state.

Transit Node

All nodes except the master node on a ring can be called transit nodes. Transit nodes monitor the status of the RRPP links that are directly connected to them, and notify the master node of the link state change. Then the master node will decide how to handle the changes.

A transit node has three states:

- Link-up state: The primary port and secondary port of the transit node are up.
- Link-down state: The primary port or secondary port of the transit node is down.
- Preforwarding state (temporarily blocked state): The primary port or secondary port of the transit node is blocked. When the link of the port of a link-down transit node goes up, the transit node changes to the preforwarding state and blocks the recovered port. When the transit node in the preforwarding state receives a packet instructing an unblock, or when the fail timer of the domain where the transit node is located expires, the transit node unblocks the blocked port.

Primary Port and Secondary Port

The master node and transit nodes all connect to an Ethernet ring through two ports. Of the two ports, one is the primary port and the other the secondary port. The port roles are user-configurable.

The primary port and secondary port of the master node function differently. The master node periodically transmits ring-check packets through its primary port. If the master node can receive the packets on its secondary port, it indicates that the RRPP ring network where the master node is located is complete. In this case, it is necessary for the master node to block its secondary port to prevent a data loop. On the contrary, if the master node does not receive the ring-check packets within a specified time, it indicates that the ring network is faulty. In this case, it is necessary for the master node to unblock the secondary port to ensure normal communication between all nodes on the ring.

The primary port and secondary port of a transit node function the same. The port roles of a transit node are also user-configurable.

In the case of a block, the secondary port on the master node of a primary ring is blocked not only from data packets but also from the protocol packets of the secondary rings. Likewise, in the case of a block, the RRPP ports (including the primary port and the secondary port) on the transit node of a primary ring are blocked from both data packets and the protocol packets of secondary rings. In the case of an unblock, these ports are opened for the packets.

RRPP Domain Timer

A domain can be configured with a domain timer; different nodes on a ring can also be configured with different domain timers. Domain timers have two types: hello timer and fail timer.

- The hello timer sets the interval for sending hello packets.
- By default, the length of a hello timer is 1s and that of a fail timer is 3s.
- The range of a hello timer is 1-10s, and the range of a fail timer is 3-30s. The configured length of the fail timer must be at least three times the length of the hello timer.

19.7.3 RRPP Packet

Packet Type

Table 19-9 lists the types of RRPP packets.

Table 19-9 Types of RRPP packets

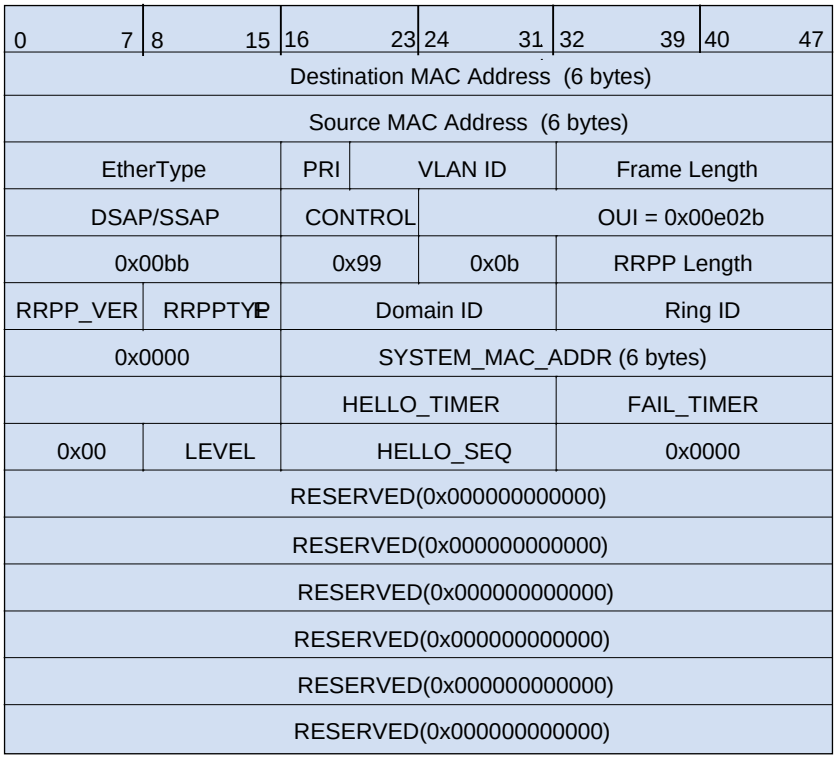
Packet Type	Description
HEALTH (HELLO)	Health-check packet. It is sent by the master node for checking the ring integrity of the network.
LINK-DOWN	Link-down packet. It is sent by the transit node, edge node, or auxiliary edge node whose direct-connection link is down. The node sends this packet to inform the master node that a link on the ring is down and the physical ring disappears.

Packet Type	Description
COMMON-FLUSH-FDB	Flush-FDB packet. It is sent by the master node to inform the transit node, edge node, or auxiliary edge node to flush their respective MAC address forwarding table.
COMPLETE-FLUSH-FDB	Ring recovery flush-FDB packet. It is sent by the master node to inform the transit node, edge node, or auxiliary edge node to flush their respective MAC address forwarding table, at the same time instructing the transit node to unblock the port that has been temporarily blocked.

RRPP Packet Format

Figure 19-45 shows the format of an RRPP packet.

Figure 19-45 RRPP packet format



The description of each field in the packet is as follows:

- Destination MAC Address: 48 bits. It indicates the destination MAC address of the packet.
- Source MAC Address: 48 bits. It indicates the source MAC address of the packet.

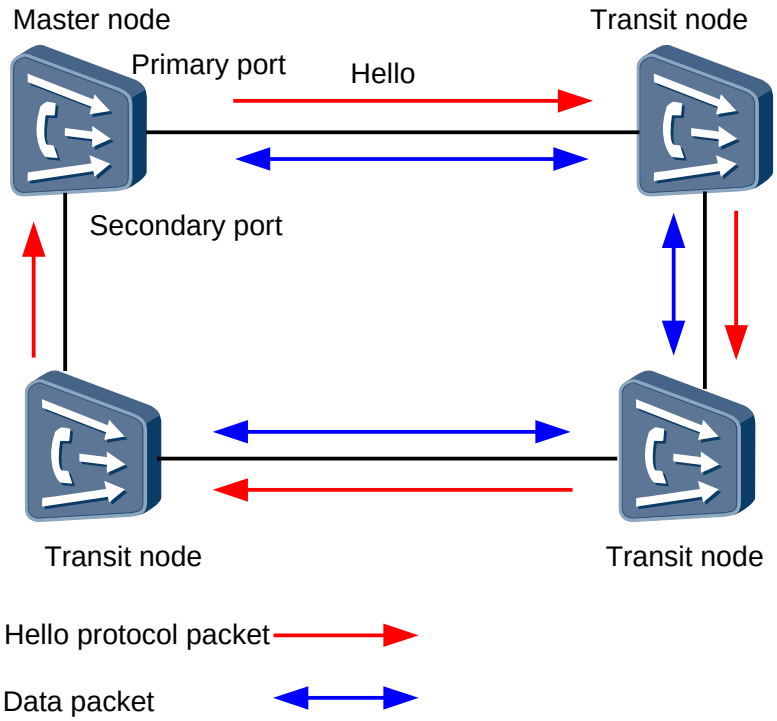
- EtherType: 8 bits. It is the packet encapsulation type field and is always 0x8100 (indicating tagged).
- PRI: 4 bits. It indicates the class of service (CoS) priority.
- VLAN ID: 12 bits. It indicates the ID of the VLAN to which the packet belongs.
- Frame Length: 16 bits. It indicates the Ethernet frame length.
- DSAP/SSAP: 16 bits. It indicates the destination service access point/source service access point.
- CONTROL: 8 bits.
- OUI: 24 bits.
- RRPP_LENGTH: 16 bits. It indicates the length of the RRPP protocol data unit.
- RRPP_VERS: 16 bits. It indicates the RRPP version.
- DOMAIN_ID: 16 bits. It indicates the ID of the RRPP domain to which the packet belongs. RING_ID: 16 bits. It indicates the ID of the RRPP ring to which the packet belongs.
- SYSTEM_MAC_ADDR: 48 bits. It indicates the bridge MAC address of the node sending the packet.
- HELLO_TIMER: 16 bits. It indicates the timeout time of the hello timer used by the node sending the RRPP packet. The timer is in the unit of second.
- FAIL_TIMER: 16 bits. It indicates the timeout time of the fail timer used by the node sending the RRPP packet. The timer is in the unit of second.
- HELLO_SEQ: 16 bits. It indicates the sequence number of the hello packet.

19.7.4 RRPP Basic Principle

Polling Mechanism

In the polling mechanism, the master node transmits the HELLO packet from its primary port periodically to check the ring network. After transmitting the HELLO packet, if the master node can receive this packet on its secondary port, it indicates that the ring network is complete. If the master node cannot receive this packet within the specified period, the master node considers that a link fault occurs on the ring network and unblocks its secondary port and allows it to forward packets. This is the basic mechanism for RRPP.

Figure 19-46 Polling mechanism implementation



The polling mechanism is a mechanism that the master node of the RRPP ring actively checks the health status of the ring network. Its process is as follows:

1. The master node transmits the HELLO packet periodically from its primary port according to the value of the HELLO timer.
2. The HELLO packet is transmitted over the ring network by passing every transit node on the ring network.
   - After transmitting the HELLO packet, if the master node can receive this packet on its secondary port before the Fail timer times out, the master node considers that the ring network is complete.
   - After transmitting the HELLO packet, if the master node cannot receive this packet on its secondary port after the Fail timer times out, the master node considers that the ring network is faulty.

After receiving the HELLO packet that is sent from the master node in the Failed state on the secondary port, the master node performs the following operations:

1. Changes itself to the Complete state.
2. Blocks its secondary port.
3. Flushes the FDB.
4. Transmits packets from its primary port to notify all the transit nodes of unblocking the temporarily blocked port and flushing their FDBs.

Mechanism of Link State Change Notification

Figure 19-47 shows the mechanism of link state change notification.

1. If a link fault occurs over the ring network, the state of the port connecting to the link is changed to Down.
2. The transit node transmits the LINK-DOWN packet actively and immediately to the master node to notify the master node of the link state change.
3. After receiving the LINK-DOWN packet, the master node considers that the ring network is faulty and unblocks its secondary port. At the same time, the master node transmits the packets to other transit nodes to notify them of flushing their FDBs.
4. After other transit nodes flush the FDB, data streams are switched to the normal links.

Figure 19-47 Link fault

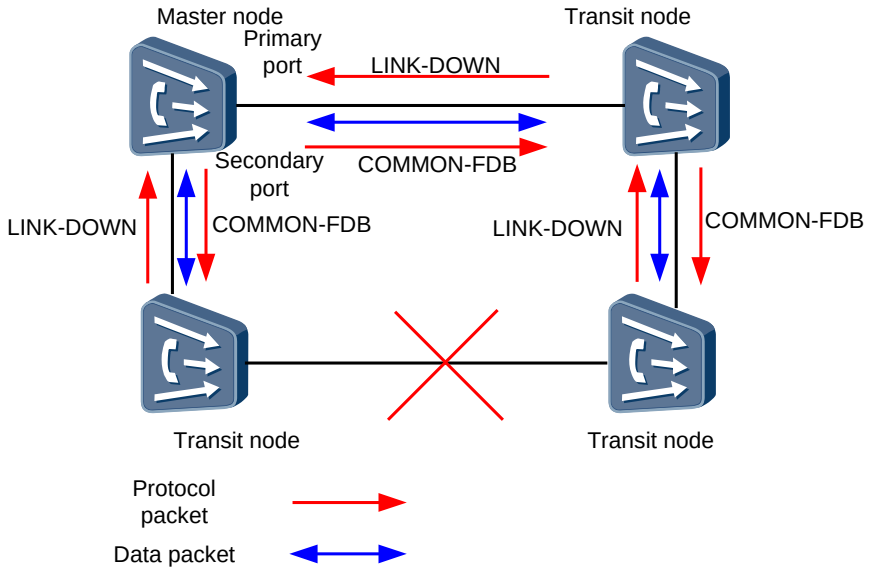
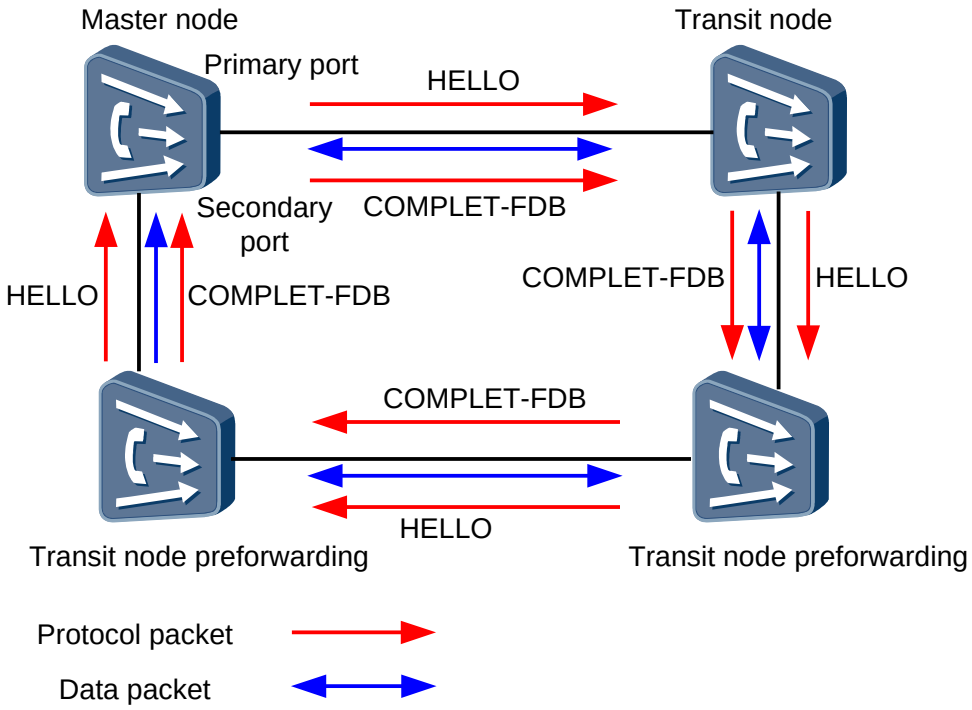


Figure 19-48 shows the mechanism of link recovery notification.

1. If a link fault is rectified, the port of the transit node changes to the Preforwarding state.
2. This transit node temporarily blocks the port whose fault is rectified; however, the HELLO packet transmitted from the master node can pass the temporarily blocked port.
3. After receiving the HELLO packet that is sent from the master node on the secondary port, the master node considers that the ring network recovers to the health state.
4. The master node blocks the secondary port and transmits the packet to other transit nodes to notify them of unblocking the temporarily blocked ports and flushing their FDBs.

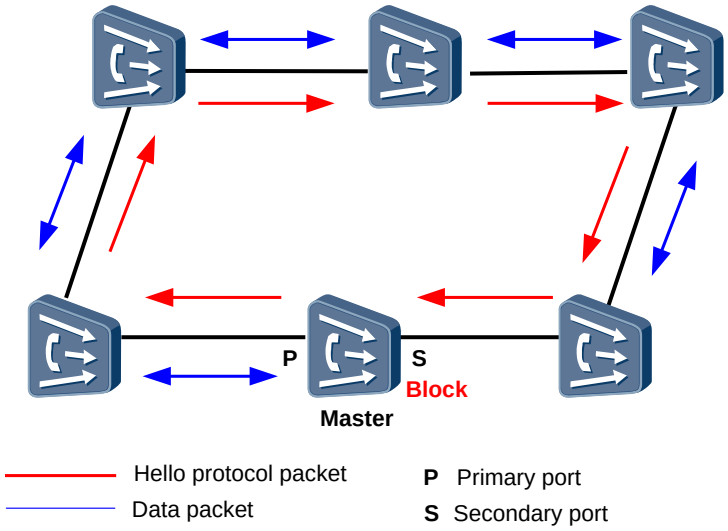
Figure 19-48 Link recovery



19.7.5 Working Principle of RRPP

Ring Polling

Figure 19-49 RRPP ring in the complete state



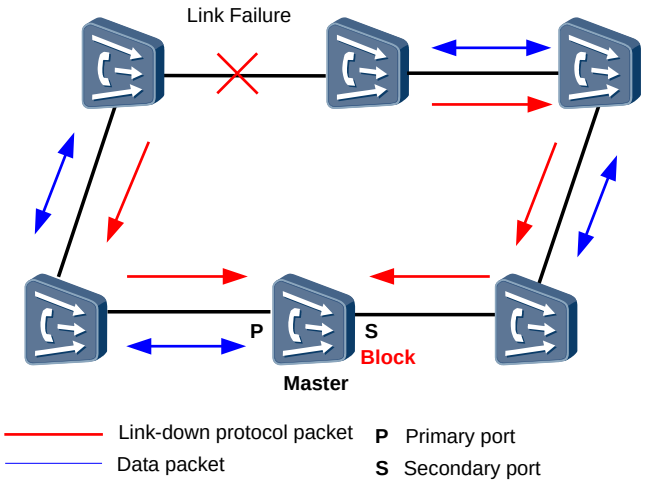
1. When all links in the entire ring network are up, the RRPP ring is in a healthy state. The state of the master node reflects the health condition of the entire ring network.

2. When the ring network is healthy, the master node needs to block its secondary port to prevent data loops. Data loops will cause a broadcast storm.
3. The master node periodically sends hello packets from its primary port. The hello packets traverse the transit nodes and finally return to the master node by its secondary port.

Link-down Alert

1. When the RRPP port of a transit node has a link-down, the transit node notifies the master node by a link-down packet, as shown in Figure 19-50.

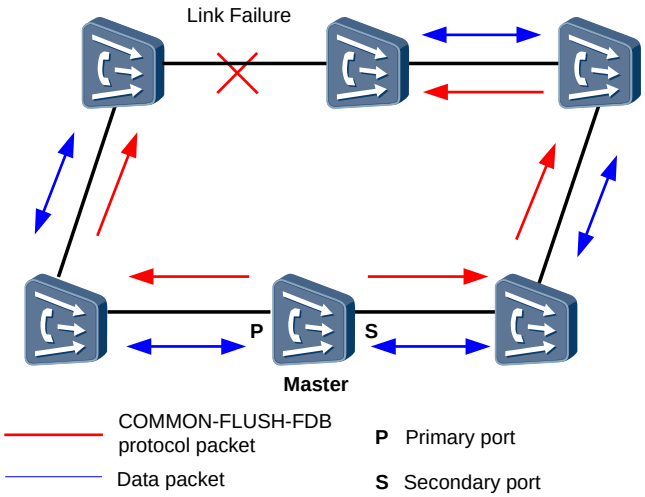
Figure 19-50 Link-down alerting of a transit node



2. After receiving the link-down packet, the master node immediately changes from the complete state to the failed state and unblocks its secondary port.

   The master node provides a polling mechanism which attends to the event that the link-down packet is lost during transmission. If the master node does not receive hello packets on its secondary port after the fail timer expires, the master node also considers that there is a ring network failure. Such a condition is processed in the same way as the transit node actively reporting link-down.
3. Since the network topology is changed, to prevent incorrect direction of packets, the master node also needs to flush its FDB table and send the COMMON-FLUSH-FDB packet from its primary port and secondary port to all transit nodes so that the transit nodes can flush their FDB tables. Figure 19-51 illustrates the process.

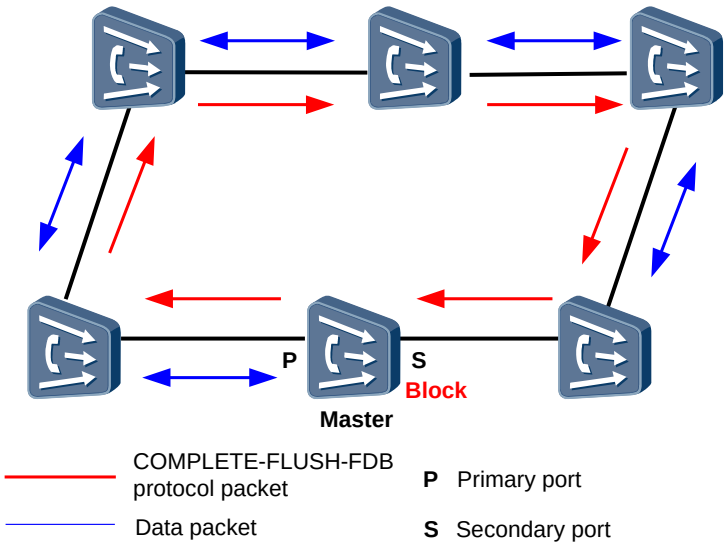
Figure 19-51 Master node changing to the failed state



Link Restoration

1. When the RRPP port of the transit node recovers, the transit node changes to the preforwarding state and blocks the recovered port.
2. The master node periodically sends hello packets from its primary port. After all the faulty links on the ring network recover, the master node will receive the hello packets on its secondary port again.
3. After the master node receives the hello packets that are sent by itself, the master node will first changes back to the complete state and block its secondary port.
4. The master node sends the COMPLETE_FLUSH_FDB packet from its primary port to notify all transit nodes to flush their FDB tables. Figure 19-52 illustrates the process.

Figure 19-52 Ring network restoration



19.7.6 RRPP Network Applications

A single-ring topology consists of only one ring. Therefore, only one RRPP domain and one RRPP ring need to be defined. The single-ring topology responds quickly in case of a network topology change and thus provides for a shorter convergence duration. This meets the requirements of a network that contains only one ring.

Normal Links

Figure 19-53 Single-ring network application (for normal links)

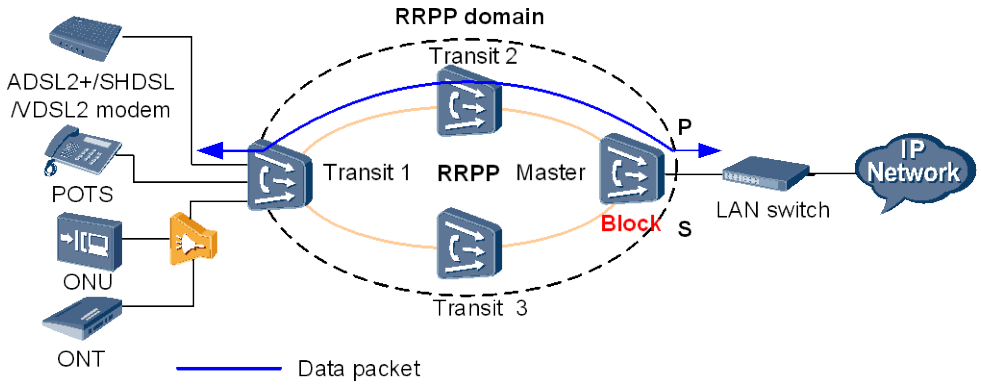


Figure 19-53 shows an RRPP single-ring topology. In normal conditions, data flows travel the "Transit 1 -> Transit 2 -> Master" route on the RRPP ring. If the link between Transit 1 and Transit 2 fails, the data flows will be rerouted on the RRPP ring.

Faulty Links

Figure 19-54 Single-ring network application (for faulty links)

As shown in Figure 19-54, when the link between Transit 1 and Transit 2 fails, the master node will receive a link-down notification and will immediately unblock its secondary port.

Now, the network topology is changed and the original MAC address tables of the nodes cannot correctly guide the forwarding any more. In this case, Layer 2 and Layer 3 service streams will be interrupted. After unblocking its secondary port, the master node immediately informs all other nodes (transit nodes) on the ring to re-learn MAC address entries and ARP entries. After entries are re-learned, Layer 2 and Layer 3 service streams on the RRPP ring will be rerouted to "Transit 1 -> Transit 3 -> Master".

19.7.7 Configuring RRPP

Rapid Ring Protection Protocol (RRPP) is a data link layer protocol specially applied to the Ethernet ring. When the Ethernet ring is complete, RRPP can prevent broadcast storms caused by a data loop. When a link on the Ethernet ring is disconnected, RRPP can quickly recover the communication channels between nodes on the Ethernet ring, increasing the network reliability.

Context

Most MANs and enterprise networks adopt the ring network structure to increase the reliability. Any faulty node on the ring does not affect the service. RRPP is a dedicated data link layer protocol applied to the Ethernet ring. Compared with other Ethernet ring technologies, RRPP has the following advantages:

- The topology convergence is quick.
- The convergence time is irrelevant with the number of nodes in the ring network. RRPP is applicable to the network that has a relatively large network diameter.
- A complete Ethernet ring can prevent broadcast storm caused by data loop.
- When a link in the Ethernet ring network is disconnected, RRPP can quickly recover the communication between nodes in the ring network by using the backup link.

Currently, the MA5600T/MA5603T/MA5608T supports only the single-ring network application of RRPP. The MA5600T/MA5603T/MA5608T can function as a primary node or a transmission node.

Procedure

- Configure the primary node.
  a. Run the **rrpp mode** command to configure the RRPP protocol mode.
     - You can select the RRPP standard mode or EAPS compatible mode. The RRPP standard mode is used by default.
     - When the RRPP function is enabled or an RRPP domain exists on the device, the RRPP protocol mode cannot be changed.
  b. Run the **rrpp domain** command to configure the RRPP domain.

     Currently, the MA5600T/MA5603T/MA5608T supports only one RRPP domain.
  c. Run the **control-vlan** command to configure the control VLAN of the RRPP domain.
     - The specified VLAN must be created through the **vlan** command and must be a standard VLAN.
     - During the configuration, you need to specify only the major control VLAN ID. The sub-control VLAN ID is specified by the system. Sub-control VLAN ID = Major control VLAN ID + 1.
     - The major control VLAN or sub-control VLAN cannot be a system reserved VLAN or a VLAN that is in use.

d. Run the **ring** command to configure the RRPP ring.
   - Currently, the MA5600T/MA5603T/MA5608T supports only one RRPP ring and the ring must be the primary ring.
   - The network role of a port joining the RRPP ring must be an upstream port. It cannot be a subtending port.

NOTE
On the same port, the RRPP function and the STP function cannot be enabled at the same time. Because the system enables the STP port-level switch by default, before creating an RRPP port, you must disable the STP function of the primary and secondary ports.

e. (Optional) Run the **timer hello-timer** command to configure the hello timer and fail time of the RRPP domain.
   - By default, the hello timer is 1s and the fail timer is 3s.
   - The value of the fail timer must be three times equal to or larger than the value of the hello timer.

f. Run the **ring enable** command to enable the RRPP ring.

g. Run the **rrpp enable** command to enable the RRPP protocol.

h. Run the **display rrpp brief domain** command to query the brief information about the RRPP domain.

i. Run the **display rrpp verbose domain** command to query details of the RRPP ring.

- Configure the transmission node.

a. Run the **rrpp mode** command to configure the RRPP protocol mode. The configuration must be the same as that on the primary node.
   - You can select the RRPP standard mode or EAPS compatible mode. The RRPP standard mode is used by default.
   - When the RRPP function is enabled or an RRPP domain exists on the device, the RRPP protocol mode cannot be changed.

b. Run the **rrpp domain** command to configure the RRPP domain. The domain ID must be the same as that on the primary node.

   Currently, the MA5600T/MA5603T/MA5608T supports only one RRPP domain.

c. Run the **control-vlan** command to configure the control VLAN of the RRPP domain. The configuration must be the same as that on the primary node.
   - The specified VLAN must be created through the **vlan** command and must be a standard VLAN.
   - During the configuration, you need to specify only the major control VLAN ID. The sub-control VLAN ID is specified by the system. Sub-control VLAN ID = Major control VLAN ID + 1.
   - The major control VLAN or sub-control VLAN cannot be a system reserved VLAN or a VLAN that is in use.

d. Run the **ring** command to configure the RRPP ring. The ring ID must be the same as that on the primary node.
   - Currently, the MA5600T/MA5603T/MA5608T supports only one RRPP ring and the ring must be the primary ring.
   - The network role of a port joining the RRPP ring must be an upstream port. It cannot be a subtending port.

NOTE
On the same port, the RRPP function and the STP function cannot be enabled at the same time. Because the system enables the STP port-level switch by default, before creating an RRPP port, you must disable the STP function of the primary and secondary ports.

e. (Optional) Run the **timer hello-timer** command to configure the hello timer and fail time of the RRPP domain.
   - By default, the hello timer is 1s and the fail timer is 3s.
   - The transmission node uses the fail timer as the timeout timer.

f. Run the **ring enable** command to enable the RRPP ring.

g. Run the **rrpp enable** command to enable the RRPP protocol.

----End

Example

To configure the MA5600T/MA5603T/MA5608T as the primary node of an RRPP ring with the following settings, do as follows:

- RRPP mode: standard
- Major control VLAN ID: 14; sub-control VLAN ID: 15
- RRPP primary port: 0/19/0; RRPP secondary port: 0/19/1
- RRPP domain ID: 1
- RRPP ring ID: 64

Other parameters adopt the default settings.

```
huawei(config)#vlan 14 standard
huawei(config)#vlan 15 standard
huawei(config)#port vlan 14-15 0/19 0-1
huawei(config)#stp port 0/19/0 disable
huawei(config)#stp port 0/19/1 disable
huawei(config)#rrpp mode rrpp
huawei(config)#rrpp domain 1
huawei(rrpp-domain-region-1)#control-vlan 14
huawei(rrpp-domain-region-1)#ring 64 node-mode master primary-port 0/19/0 second
ary-port 0/19/1 level 0
huawei(rrpp-domain-region-1)#ring 64 enable
huawei(rrpp-domain-region-1)#quit
huawei(config)#rrpp enable
huawei(config)#display rrpp brief domain 1
  ------------------------------------------------------------------------------
  Rrpp Protocol Status  : Enable
  Rrpp protocol mode    : RRPP
  Number of RRPP Domains: 1
  ------------------------------------------------------------------------------
  Domain Index        : 1
  Major Control VLAN    : 14
  Hello Timer         : 1  sec (default is 1 sec)
  Fail Timer          : 3  sec (default is 3 sec)
  Number of RRPP Rings  : 1
  ------------------------------------------------------------------------------
   Ring   Ring    Node    Primary/Common      Secondary/Edge      Is
   ID     Level   Mode    Port               Port               Enabled
   ------------------------------------------------------------------------------
   64     0       M       GE 0/19/0          GE 0/19/1          Yes
  ------------------------------------------------------------------------------
  Note: M - Master, T - Transit , E - Edge , A - Assistant-Edge
```

19.7.8 RRPP Reference Standards and Protocols

The following lists the reference standards and protocols of this feature:

- RRPP (by Huawei)
- Ethernet Automatic Protection Switching (EAPS)

19.8 ERPS

Ethernet ring protection switching (ERPS) is a ring network protocol defined in Recommendation ITU-T G.8032. Interoperation is achieved if all devices participating in the ring network support ERPS. On a network composed of ERPS-supporting devices, the service protection solution with quick switching can be implemented using a small number of links.

19.8.1 Introduction to ERPS

To resolve loop issues, Huawei has supported equipment supporting the Spanning Tree Protocol (STP) and Rapid Ring Protection Protocol (RRPP). The following challenges need to be addressed:

- As higher requirements are posed on the switching time for Layer 2 Ethernet, STP supports link recovery on a basis of seconds, falling short of the carrier-class convergence performance requirements.
- RRPP is a proprietary protocol. Interoperation is not supported if a ring network is composed of devices of multiple vendors.

Ethernet ring protection switching (ERPS) is a ring network protocol defined in Recommendation ITU-T G.8032. Interoperation is achieved if all devices participating in the ring network support ERPS. In addition, ERPS supports fast convergence, meeting carrier-class reliability requirements. Currently, ERPS has two versions, V1 and V2.

- V1 supports only single-ring protection, single Ethernet ring protection (ERP) instance, and revertive switching.
- Based on V1, V2 supports multiple-ring protection, multiple ERP instances, and non-revertive switching. V2 also optimizes the flushing mechanism for the filtering database (FDB), and supports commands for a forced switchover, manual switchover, or a clearing operation.

The MA5600T/MA5603T/MA5608T implements functions of the V1 protocol using the V2 state machine, supports V1 and V2 message reception and V1 message transmission, and supports all functions of V1 and the multiple instances defined by V2.

Recommendation ITU-T G.8032 only protects MAC address+VLAN services; the MA5600T/MA5603T/MA5608T extends the protocol by protecting multicast and Layer 3 services.

The following explains the concepts involved in the description above.

Concept	Description
Single ring	An ERPS ring physically composed of a set of devices that are connected to form a closed loop.
Multiple rings	A complex network composed of multiple interconnected single rings. The single rings may be interconnected through an interconnection

Concept	Description
	point or through dual interconnection nodes.
Single instance	A mechanism that allows only one logical ERPS ring to run on a physical ring network. All services are protected by this ring.
Multiple instances	A mechanism that implements protection switching for ERPS rings based on Multiple Spanning Tree Protocol (MSTP) instances. Multiple logical ERPS rings can run on one physical ring network. The topologies of the rings are independent of each other. These rings implement ring automatic protection switching (R-APS) protocol message exchange, fault processing, and service switchover. Therefore, each of the rings must protect a specific type of service. When the network runs in normal conditions, each ERPS ring can select its own blocking point to load balance service traffic over the rings. The selection can be implemented through configuration.
Revertive switching	A mechanism that blocks the ring protection link (RPL) and reverts to the original faulty link to carry service traffic when the fault causing the switch is cleared and the wait-to-restore (WTR) timer expires.
Non-revertive switching	A mechanism that still blocks the original faulty link after the fault causing the switch is cleared.
Forced switching	A port configured with forced switching is immediately blocked, regardless of whether other links on the ring have encountered a fault.
Manual switching	A port configured with manual switching is blocked only when the ring does not contain a faulty link or a port configured with forced switching.
clear command	A command that clears the forced switching and manual switching locally configured.

19.8.2 Basic Concepts of ERPS

This topic describes basic concepts related to the Ethernet ring protection switching (ERPS) feature to help you understand the working principles of the ERPS ring.

ERPS Ring

An ERPS ring is composed of a group of interconnected Layer 2 switches that are configured with the same control VLAN. an ERPS ring is the basic element for implementing the ITU-T G.8032 protocol. A Layer 2 switch participating in the ERPS ring is called a node. On the node, the ports that participate in the ERPS ring are called ring ports.

Port Roles

The ITU-T G.8032 protocol defines two port roles: ring protection link (RPL) owner port and common port. The link where an RPL owner port resides is called a ring protection link. The device where an RPL owner port resides is called an RPL owner node.

- **RPL owner port**

An ERPS ring has only one RPL owner port, which can be configured by users. Loop avoidance is achieved on an ERPS ring by blocking the RPL owner port. A blocked RPL owner port cannot forward data but can send ring automatic protection switching (R-APS) messages.

As defined in the IEEE 802.1ag protocol, the down maintenance end point (MEP) on the RPL owner port can receive and send Ethernet OAM messages.

When an RPL owner node receives a fault message and learns that a node or link on the ERPS ring is faulty, the RPL owner node automatically enables the RPL owner port. Then, the RPL owner port resumes traffic reception and transmission, ensuring uninterrupted traffic.

NOTE

For load balancing purposes, it is recommended to select a device with the following characteristics as an RPL owner node: the device is located close to the user side; an equivalent (or approximately so) number of nodes are located on the link between each of the two ring ports of the device and the convergence node.

- **Common port**

  On an ERPS ring, all ports except the RPL owner port are common ports. Common ports are responsible for monitoring the status of their directly connected links, and notifying any link status changes to the ports on other nodes.

Port States

On an ERPS ring, a port has two states:

- **Forwarding**

  In forwarding state, a port can forward data, and receive and send R-APS messages.

- **Discarding**

  In discarding state, a port cannot forward data, but can receive and send R-APS, EFM, CFM, LLDP, and 802.1x messages.

R-APS Messages

Defined in the ITU-T G.8032 protocol, R-APS messages are protocol messages that run on ERPS rings to notify nodes of fault occurrence or fault clearing. R-APS messages include the following types:

- **R-APS signal fail (SF)**

  indicates a link failure. When detecting a link failure, a node sends an R-APS (SF) message to notify other nodes on the ring. When receiving an R-APS (SF), the RPL owner node unblocks the RPL owner port so that the RPL owner port forwards data. Other nodes determine whether to flush the FDT according to the DNF flag in the R-APS (SF) message.

- **R-APS no request (NR)**

  indicates a link recovery. When detecting a link recovery, a node sends an R-APS (NR) message to notify other nodes on the ring. When receiving an R-APS (NR), the RPL owner node starts the WTR timer. Other nodes determine whether to flush the FDT according to the DNF flag in the R-APS (NR) message.

- **R-APS no request, RPL blocked (NR, RB)**

  An R-APS (NR, RB) message is a type of R-APS (NR). It is sent by an RPL owner to signal that the network is normal and the RPL has been blocked. When receiving an R-APS (NR, RB), other nodes on the ring will unblock the common ports, and determine whether to flush the FDT according to the DNF flag in the message.

The following explains the acronyms involved in the description above.

Acronym	Full Name	Meaning
NR	No request	The network is in normal conditions.
RB	RPL blocked	An RPL is blocked.
DNF	Do not flush	A flag in the R-APS message fields and indicates whether to flush the FDB.
FDB	Filtering database	A forwarding database on a node.
SF	Signal fail	The network encounters a fault.

Timers

The ITU-T G.8032 protocol defines the following timers for an ERPS ring.

- **Guard timer**

  After a node on a Layer 2 network running ERPS recovers from a link fault or node fault, the node sends an R-APS (NR) message to other nodes on the ring and starts the guard timer. The node does not receive R-APS messages before the timer expires, which prevents the node from unblocking a recovered port after receiving expired R-APS(SF) messages. The unblocking will cause a loop on the network. If the node receives R-APS (SF) messages after the timer expires, the node changes its port state to forwarding.

- **Holdoff timer**

  Layer 2 networks running ERPS require different protection switching sequences. For example, when a node becomes faulty in multiple-layer service application, carriers prefer a duration to rectify the fault and the duration is not perceived by subscribers. To meet the carrier requirement, protection switching is not performed immediately after the fault occurs. The fault is reported if it is not rectified after the holdoff timer expires.

- **WTR timer**

  On a Layer 2 network running ERPS, an RPL owner port enters the forwarding state when a node or link on the ERPS ring fails. When the node or link recovers, some ports on the node or link may still stay in the Down state (that is, not Up yet). If the RPL owner port is immediately blocked upon the fault recovery, and finds that there are still ports in the Down state, the RPL owner port will be unblocked again. To avoid such toggling protection states in the case of intermittent faults, the RPL owner port starts a wait-to-restore (WTR) timer when receiving the R-APS (NR) message from the other ports.

  - If the RPL owner port receives the R-APS (SF) message from the other ports before the WTR timer times out, the RPL owner port stops the WTR timer and stays in the forwarding state.
  - If the RPL owner port does not receive the R-APS (SF) message when the WTR timer times out, the RPL owner port is blocked and sends the R-APS (NR, RB) message to other ports. When receiving the R-APS (NR, RB) message, the other ports switch to the forwarding state.

19.8.3 ERPS Principles

This topic uses ERPS network diagrams to describe the fault detection and recovery processes of an ERPS ring.

Network Diagram

Figure 19-55 shows the network diagram of an ERPS single ring. The nodes on the ERPS ring exchange R-APS messages to notify each other of the port status changes and coordinate to implement protection switching for services. The port status changes on a ring node depend on detection of local port status, Ethernet OAM detection results, and the port status changes on other nodes of the ring.

Figure 19-55 Network diagram of an ERPS single ring

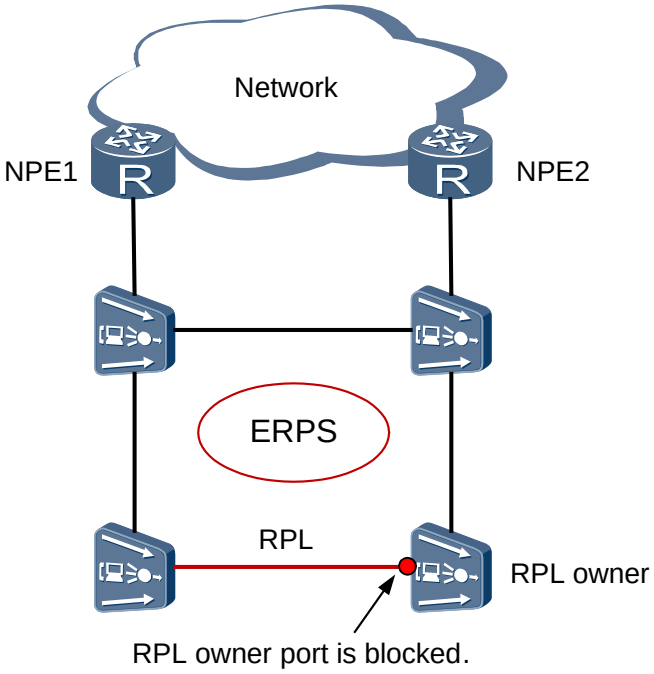
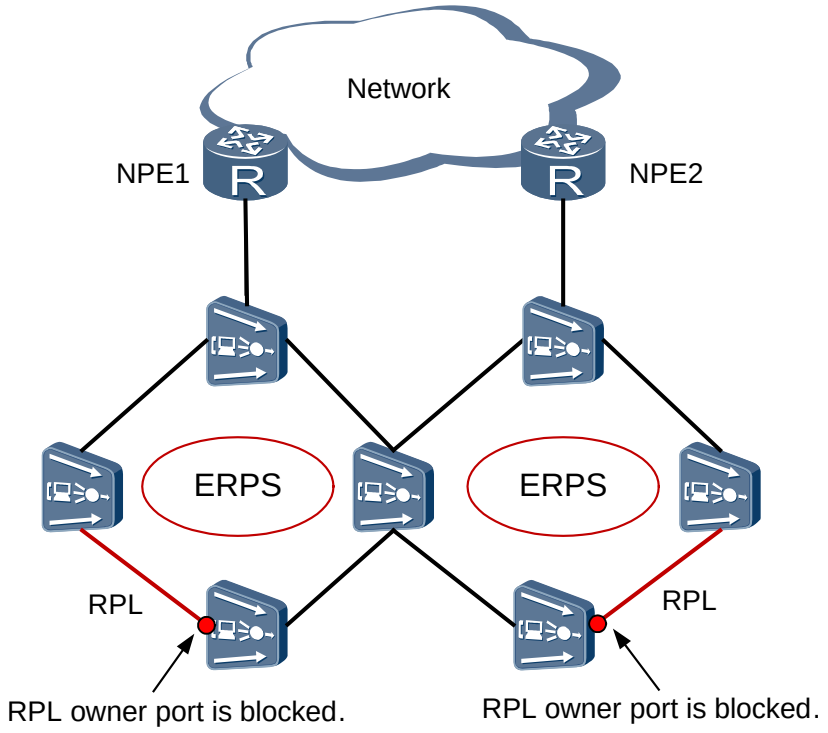


Figure 19-56 shows the network diagram of ERPS interconnected rings. The two physical ring networks are interconnected through one node. Four ports on the interconnection node are configured with ERPS. Two of the four ports locate on one ring, and the other two on the other ring. Except the interconnection node, other nodes on the two interconnected rings function in the same mechanism as the nodes on a single ring. The R-APS protocol messages of the two rings are independent of each other, so the two rings need different control VLANs.

The interconnection node can be configured with multiple ERPS rings, and the rings can have different numbers of nodes.

Figure 19-56 ERPS rings interconnected through one node



Configuration Items of ERPS Ring

Each ring node must be configured with the following items:

- Control VLAN

  The control VLAN transmits ERPS messages and does not forward data messages, which improves ERPS security.

- Protection instance

  Protection instances are implemented by configuring mappings between MSTP instances and VLANs.On a Layer 2 network with ERPS enabled, the VLAN that transmits ERPS messages and data messages must be mapped to a protection instance so the ERPS ring forwards or blocks the messages based on the blocking principles. Without proper forwarding or blocking, the messages may cause a broadcast storm on the ring network and render the network unavailable.

- Port role

  Configuration related to port roles is to add ports to an ERPS ring, and specify the RPL owner port on the RPL owner node.

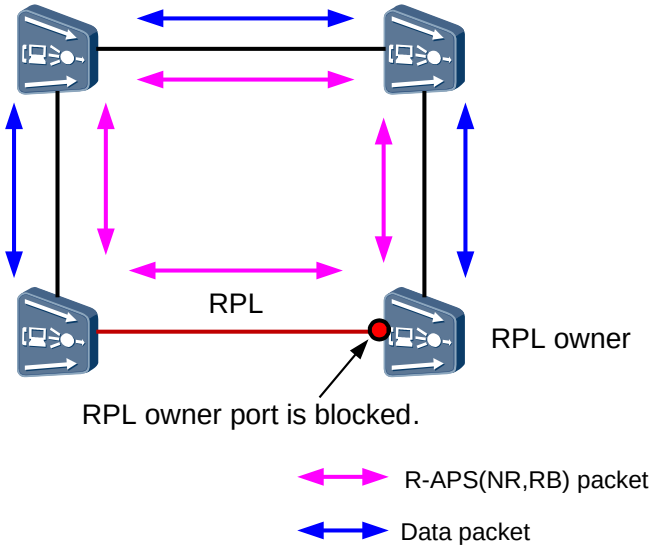
Fault Detection and Recovery

The following users a single ERPS ring as an example to describe the fault detection and recovery processes of an ERPS ring.

In normal conditions, all links on the network are in the UP state, and the RPL owner port is blocked, therefore forming link redundancy backup and avoiding loops. Service traffic is forwarded through all links except the RPL. In this situation, the R-APS messages travelling

on the network are all R-APS (NR, RB) messages that are sent by the RPL owner node, as shown in Figure 19-57.

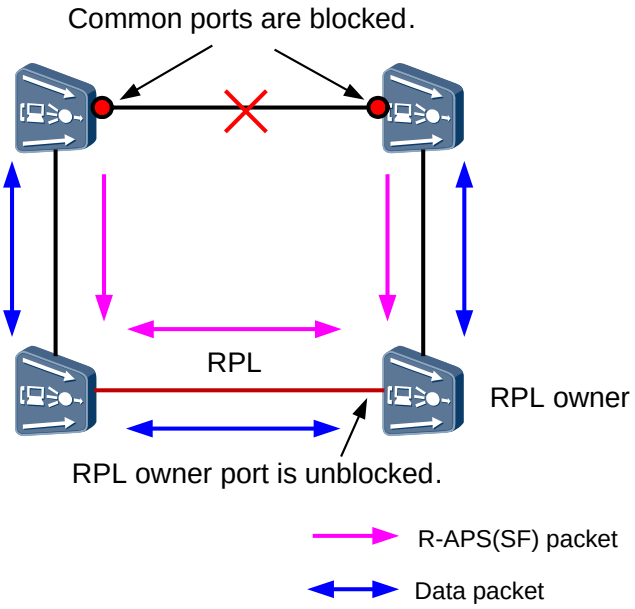
Figure 19-57 An ERPS ring network in the normal state



When a fault occurs on a link, the nodes connected to the faulty link detect the fault and start the holdoff timer. If the fault persists after the holdoff timer times out, the nodes block the ports on the faulty link, and send R-APS (SF) messages to other nodes on the ring. After receiving the R-APS (SF) message, the RPL owner node unblocks the RPL owner port, and switches to the RPL to transmit service traffic. In this way, ring network protection is implemented. When receiving the R-APS (SF) message for the first time, the node will flush its FDB, and the ring network enters the protection state, as shown in Figure 19-58. The node will not trigger any operation when receiving subsequent R-APS (SF) messages.

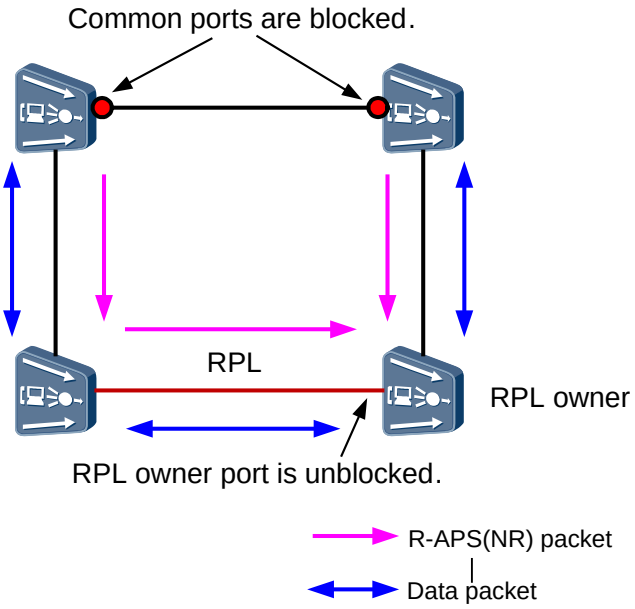
To implement fast switching, each node detecting the link fault will rapidly send 3 R-APS (SF) messages, at a 3.33 ms interval, which equals the interval of fast detection by using CCM messages. After sending the 3 R-APS (SF) messages, the node will send one R-APS (SF) message every 5s. Such a mechanism ensures fast network switching without burdening the devices.

Figure 19-58 An ERPS ring network in the protection state



When the fault recovers on the link, the nodes connected to the originally faulty link detect the fault recovery, start the guard timer, and stay in the blocked state. At the same time, the nodes send R-APS (NR) messages to the RPL owner node, notifying the RPL owner node that the RPL owner port can be blocked, as shown in Figure 19-59.

Figure 19-59 ERPS ring network status after fault recovers and before WTR timer times out



After receiving the R-APS (NR) message, the RPL owner node blocks the RPL owner port only after the WTR timer expires. Meanwhile, the RPL owner node sends the R-APS (NR, RB) message. The other nodes on the ring flush their FDB when receiving the R-APS (NR, RB) message. After the guard timer times out, the node that detects the fault recovery stops

periodically sending R-APS (NR) messages and unblocks the local ports. When the network resumes the normal state, the RPL owner node continues to send R-APS (NR, RB) messages.

19.8.4 Configuring ERPS

The device supports ERPS as defined in ITU-T G.8032. Interoperation is achieved if all devices participating in the ring network support ERPS. The ERPS feature enables fast convergence and meets carrier-class reliability requirements.

Prerequisites

The service VLANs to be protected have been configured.

Context

- In the single ERPS ring scenario, the ring must be configured on all nodes participating in the ring. In the multiple-ERPS-ring scenario, each ring has its own configurations.
- Note that the system does not have a restriction mechanism for some configuration items. When planning and configuring such items, make sure that they are planned and configured correctly, in line with system requirements. The following table provides the configuration restrictions on ERPS.

Item	Limitation	Does the Device Check the Limitation?
Control VLAN	• A control VLAN must be a standard VLAN and cannot be used by a Rapid Ring Protection Protocol (RRPP) ring or a smart link. • A control VLAN cannot be a reserved VLAN. • An ERPS ring cannot be modified or deleted if a port is added to the ring. To modify or delete a control VLAN, delete the port from the ERPS ring first. • Different ERPS rings must be configured with control VLANs of different IDs on the same node.	Yes
	• The control VLAN must be different from the service VLAN. • A control VLAN cannot be configured as a native VLAN. • The attribute of a control VLAN must be common. • All the devices in an ERPS ring must be configured with the same control VLAN.	No. The user must correctly configure this item.
Ring port	• Each node supports up to 2 ports to join the same ERPS ring. • MSTP, RRPP, smart link, Ethernet port protection, and ERPS are mutually exclusive. • Upstream ports and cascade ports can function as ring ports, but user-side ports cannot.	Yes

Item	Limitation	Does the Device Check the Limitation?
	• When a control VLAN is created for an ERPS ring or a port is added to an ERPS ring, the port is not added to the control VLAN forcibly. Ensure that all the ports added to the ERPS ring are in the control VLAN. If a port is not in the control VLAN, the ERPS packets sent to the port are discarded. • A ring port must not be configured with static MAC addresses or static ARP entries. • On all nodes except the convergence node, it is not allowed to isolate two ring ports using the port isolation function. • On all nodes, ensure that packets of the control VLAN can be transmitted and received between two ring ports of the same node. • The ring ports on the interconnection node of an access-layer ERPS ring and a convergence network can be configured as cascade ports or upstream ports, but the ring ports on other nodes on the rings can be configured as upstream ports only. • In a network where an access-layer ERPS ring is connected to the upstream network through multiple nodes, the ring ports on all nodes of the ERPS ring must be upstream ports. • The interconnection node of an access-layer ERPS ring and a convergence network does not support associated routes or Neighbor Discovery (ND) entries. If the interconnection node is enabled to automatically generate associated routes and ND entries, when the topology of the ERPS ring changes, the DHCPv6 users under all nodes on the ring must dial up again. Otherwise, services may fail.	No. The user must correctly configure this item.
Protection instance	• When the SCUB control board is used, a VLAN whose TAG Protocol ID (TPID) has been changed cannot be mapped to a protection instance. • A protection instance cannot belong to multiple ERPS rings at the same time. • One ERPS ring can be configured with one or multiple protection instances. Each protection instance maps one or multiple VLAN services.	No. The user must correctly configure this item.

Procedure

Run the **vlan** command to create a control VLAN.

Step 1 Run the **port vlan** command to add ERPS ring ports to the control VLAN.

Step 2 Map the control VLAN and service VLANs to the protection instances.

1. Run the **stp region-configuration** command to enter the MSTP region mode.
2. Run the **instance vlan** command to map the service VLANs and the control VLAN to the specified MSTP instance.
3. Run the **active region-configuration** command to activate the configuration of the MSTP region.
4. Run the **quit** command to quit the MSTP region mode.

Step 3 Run the **erps ring** to create an ERPS ring and enter the ring mode.

The nodes on the same ring must be configured with the same ring ID.

Step 4 Run the **control-vlan** command to configure the control VLAN for the ERPS ring.

vlan-id must be the same as the *vlan-id* specified in the **vlan** commamd.

Step 5 Run the **protected-instance** command to configure the protection instance for the ERPS ring.

instance-id must be the same as the *instance-id* specified in the **instance vlan** commamd.

Step 6 Run the **port** command to add the ports to the ERPS ring.

port-id must be the same as the *port-id* specified in the **port vlan** commamd. Specify the **rpl owner** parameter when configuring the RPL owner port.

Step 7 (Optional) Configure ERPS parameters.

NOTE
It is recommended to configure the same ERPS parameters on all nodes of the ring.

- Configure the timers.
  - Run the **guard-timer** command to configure the guard timer started by a node when the node detects a fault recovery.
  - Run the **holdoff-timer** command to configure the holdoff timer started by a node when the node detects a fault.
  - Run the **wtr-timer** command to configure the WTR timer started by a node when the node receives an R-APS (NR) message.
- Run the **description** command to configure the description of the ERPS ring.
- Run the **priority** command to set the 802.1p priority of ERPS messages on the ERPS ring.
- Run the **raps-mel** command to set the value for the MEL field in the ring automatic protection switching (RAPS) messages on the ERPS ring.

Step 8 Run the **display erps** command to query the ERPS configuration.

----End

Example

The following is an example of configuring the RPL owner node by using the following ERPS parameters:

- ERPS ring ID: 1
- Control VLAN ID: 20
- Protection instance ID: 3; mapping service VLAN IDs: IDs 2-10, 12-16
- Ring ports: 0/19/0 and 0/19/1, of which 0/19/0 is the RPL owner port
- ERPS ring description: ERPS Ring 1 The other parameters use their default values.

```
huawei(config)#vlan 20 standard
huawei(config)#port vlan 20 0/19 0,1
huawei(config)#stp region-configuration
huawei(stp-region-configuration)#instance 3 vlan 2 to 10 12 to 16 20
huawei(stp-region-configuration)#active region-configuration
huawei(stp-region-configuration)#quit
huawei(config)#erps ring 1
huawei(config-erps-ring1)#control-vlan 20
huawei(config-erps-ring1)#protected-instance 3
huawei(config-erps-ring1)#port 0/19/0 rpl owner
huawei(config-erps-ring1)#port 0/19/1
huawei(config-erps-ring1)#description ERPS Ring 1
huawei(config-erps-ring1)#quit
huawei(config)#display erps
 ------------------------------------------------------------------------------
 Ring  Control  WTR Timer  Guard Timer  Port 1         Port 2
 ID    VLAN     (min)      (csec)
 ------------------------------------------------------------------------------
   1      20         5         200  (D,R)0/19/0     (D)0/19/1
 ------------------------------------------------------------------------------
 D : Discarding, F : Forwarding, R : RPL Owner
 Total number of rings configured : 1
```

19.8.5 ERPS Reference Standards and Protocols

The reference standards and protocols of the ERPS feature are as follows:

- ITU-T G.8032 V1
- ITU-T G.8032 V2

19.9 STM-1 Port Protection Switching

This topic describes the feature of STM-1 port protection switching.

19.9.1 Introduction to STM-1 Port Protection Switching

Definition

The MA5600T/MA5603T/MA5608T supports TDMoGEM and SAToP (TDM PWE3). The MA5600T/MA5603T/MA5608T can terminate the TDM service on its STM-1 port to

interoperate with other SDH devices. To ensure the reliability of lines, STM-1 port protection is required. The MA5600T/MA5603T/MA5608T supports board-level protection (1+1 protection) between two boards and port-level protection within a board. With this feature, the MA5600T/MA5603T/MA5608T can automatically cope with network faults.

NOTE
STM-1 port protection is not applicable to a ring network.

Purpose

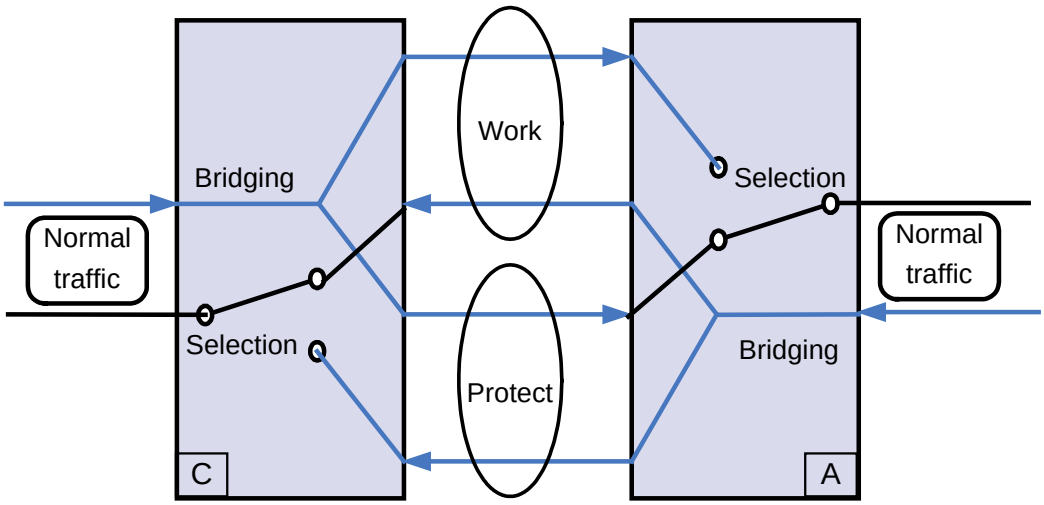
When a hardware fault or line fault occurs on the STM-1 port under protection, the system immediately switches the service from the faulty port to the backup functional unit.

19.9.2 STM-1 Port Protection Switching Principle

The fundamental purpose of STM-1 port protection is to quickly switch services to the line of another STM-1 port in case of the failure of one STM-1 port, thus implementing quick switching on the physical plane. If the line status of the working port becomes abnormal, the system should be notified immediately to switch services to the backup member port.

Figure 19-60 illustrates the principle of the 1+1 line unidirectional protection of STM-1 ports.

Figure 19-60 Principles of the 1+1 line unidirectional protection of STM-1 ports



NOTE
In Figure 19-60, C and A respectively represent the two devices connected to the STM-1 ports, **Work** indicates that the corresponding STM-1 port is the working port, and **Protect** the protection port.

Figure 19-60 shows that protection switching is performed in one direction. When traffic is transmitted from device C to device A, device C transmits signals to active and standby channels at the same time and device A determines which port is to receive the signals according to the running conditions of the ports. After device A selects the port to receive signals, the port receives signals through a selection switch. Then, the line of the port that receives signals changes to the working state automatically.

If the member ports of a protection group are on different boards, the service packets from the UNI side are copied by the LAN switch of the control board and are transmitted to the board where active and standby STM-1 ports are located for processing. In the upstream direction,

active and standby STM-1 ports transmit the same traffic. In the downstream direction, the active port is selected to receive traffic and to forward traffic to the UNI side; the traffic received by the standby port is dropped.

If the member ports of a protection group are on the same board, the service packets from the UNI side are copied by the logic chip of the daughter board to the transmit direction of the standby STM-1 port; in the downstream direction, the active port is selected to receive traffic and to forward traffic to the UNI side.

The system automatically selects active and standby ports according to the physical status and line status of the ports, and does not use the APS protocol to determine protection switching.

19.9.3 Configuring the MPLS Service Board Redundancy Backup

This topic describes how to configure 1+1 redundancy backup for the MPLS service board. In this way, when the MPLS service board is faulty, the service is not affected.

Context

Only MPLS boards of the same type support redundancy backup.

Procedure

Create a protection group.

Run the **protect-group** command to a protection group that protects the service processing board.

- Configure **protect-target** to **service-process-board**.
- The working mode of the MPLS service board protection group can be only **boardstate**.

Step 1 Add members to the protection group.

Run the **protect-group member** command to add members to a protection group.

- When adding members to the protection group, add a working member, and then add a protection member.
- Adding a protection group member based on the port is not supported for the MPLS service board, and only adding a protection group member based on the board is supported.

Step 2 Enable the protection group.

Run the **protect-group enable** command to enable the protection group. After a protection group is created, the protection group is in the disabled state by default. You should enable the protection group to make the configuration take effect.

Step 3 Query the information about the protection group.

Run the **display protect-group** command to query the information about the protection group and all the members in the protection group.

----End

Example

To configure redundancy back for MPLS boards in slots 0/2 and 0/3 of the MA5600T/MA5603T/MA5608T so that when the service board in slot 0/2 fails, the system can automatically switch the services to the service board in slot 0/3.

```
huawei(config)#protect-group 1 protect-target service-process-board workmode
boardstate
huawei(protect-group-1)#protect-group member board 0/2 role work
huawei(protect-group-1)#protect-group member board 0/3 role protect
huawei(protect-group-1)#protect-group enable
```

19.9.4 STM-1 Port Protection Switching Reference Standards and Protocols

The following lists the reference standard of this feature:

- ITU-T G.841

19.10 BFD

19.10.1 Overview

Purpose

Bidirectional forwarding detection (BFD) rapidly monitors communications faults between systems and notifies upper-layer applications of those faults.

Description

To minimize the impact of a fault on services and improve network availability, a network device must rapidly detect communications faults between adjacent devices so that the upper layer protocol can resolve the issue and recover services.

Currently, the existing detection mechanisms are as follows:

- Hardware detection: For example, Synchronous Digital Hierarchy (SDH) alarms are used to detect link faults. Hardware detection can fast detect a fault; however, not all media support this hardware detection mechanism.
- Slow Hello: Usually refers to the Hello mechanism used by a routing protocol. The slow Hello mechanism can detect a fault in seconds. For example, in high-speed gigabit rate data transmission, a detection time of more than one second results in a large data loss. Delay-sensitive services, like voice, cannot function with more than a one second delay.
- Other detection mechanisms: Different protocols or manufacturers may provide their own proprietary detection mechanisms; however, deploying proprietary detection mechanisms on different systems can be very difficult.

BFD has been developed to supplement other detection mechanisms.

BFD provides the following features:

- Low-cost fast fault detection for channels between adjacent forwarding engines. Faults can be detected on interfaces, data links, and forwarding engines.
- A single mechanism capable of real-time detection over any media, at any protocol layer.

19.10.2 Key Concepts

BFD detects communications faults between forwarding engines, specifically the connectivity of a data protocol on a path between systems. The path can be a physical link, a logical link, or a tunnel.

BFD can be regarded as a service provided by the system.

- Upper layer applications provide BFD with parameters, such as the detection address and the detection time.
- BFD creates, deletes, or modifies a BFD session according to this information and notifies the upper layer applications of the session status.

BFD offers the following features:

- Low-cost, fast detection of path faults between adjacent forwarding engines
- A single mechanism capable of detection over any media, at any protocol layer, facilitating an integrated detection mechanism.

The following sections describe basic BFD concepts, including the BFD detection mechanism, detected link types, BFD session modes, and session management.

BFD Detection Mechanism

In the BFD detection mechanism, two systems set up a BFD session, and periodically send BFD control packets along the path between them. If one system does not receive BFD control packets within a specified period, the system concludes that a fault has occurred on the path.

BFD control packets are encapsulated in UDP packets. In the initial phase of a BFD session, both systems negotiate with each other using parameters in BFD control packets, such as discriminators, expected minimum intervals for sending and receiving BFD control packets, and local BFD session status. When negotiations are successful, the two systems send BFD control packets to each other at the negotiated intervals.

To meet fast detection requirements, the BFD draft specified that BFD control packets must be sent and received at intervals expressed in microseconds. However, BFD-enabled devices of most manufacturers can only process BFD control packets within milliseconds due to limited processing capabilities. Therefore, the configured interval is expressed in milliseconds and is converted to microseconds during internal processing.

BFD provides the following detection modes:

- Asynchronous mode: The main mode is asynchronous mode. In asynchronous mode, two systems periodically send BFD control packets to each other. If one system fails to receive packets consecutively, the BFD session is considered Down.
- Query mode: The second mode is the query mode. If multiple BFD sessions exist in a system, periodically sending BFD control packets can draw significant system resources. To prevent this, you can use the query mode. In query mode, after a BFD session is set up, the system does not periodically send BFD control packets, but detects the connectivity through another mechanism (such as the Hello mechanism of a routing protocol or the hardware detection mechanism), reducing system resources used by the BFD session.

An auxiliary function of the two modes is the Echo function. When the Echo function is activated, a BFD control packet is sent as follows: The local system sends a BFD control packet and the remote system sends the BFD control packet back through the forwarding channel. If consecutive Echo packets are not received, the BFD session is declared Down. The Echo function can work in asynchronous or query mode.

At present, only the passive Echo function is supported.

Types of Links That Can Be Detected by BFD

Table 19-10 Types of links detected by BFD

Link Type	Classification	Description
IP links	• Layer 3 physical interfaces • Ethernet sub-interfaces (including Eth-Trunk sub-interfaces)	If a physical Ethernet interface has multiple sub-interfaces, BFD sessions can be separately established on the physical Ethernet interface and its sub-interfaces.
IP-Trunks	• IP-Trunk links • IP-Trunk member links	Separate BFD sessions can be established to detect link faults on an IP-Trunk and its member interfaces at the same time.
Eth-Trunks	• Layer 2 Eth-Trunk links • Layer 2 Eth-Trunk member links • Layer 3 Eth-Trunk links • Layer 3 Eth-Trunk member links	Separate BFD sessions can be established to detect link faults on an Eth-Trunk and its member interfaces at the same time.
VLANIF	• VLAN Ethernet member links • VLANIF interfaces	Separate BFD sessions can be established to detect link faults on a VLANIF interface and its member interfaces at the same time.
MPLS LSPs	• In static mode, BFD can detect the following types of LSPs: – Static LSPs – LDP LSPs – TE tunnels, static CR-LSPs bound to tunnels, and RSVP CR-LSPs • In dynamic mode, BFD can detect the following types of LSPs:	• BFD can detect a TE tunnel that uses CR-Static or RSVP-TE as its signaling protocol and detect the primary LSP bound to the TE tunnel. • A dynamic BFD session cannot detect the entire TE tunnel.

Link Type	Classification	Description
	– LDP LSPs – Static CR-LSPs bound to tunnels and RSVP CR-LSPs – LDP Tunnel	
PWs	• SS PWs • MS PWs • BGP PWs	-

BFD Session Modes

A BFD session can be set up in the following modes:

BFD differentiates sessions by My Discriminator and Your Discriminator in the control packets. The main difference in establishment of static and dynamic BFD sessions is that My Discriminator and Your Discriminator are set differently.

Table 19-11 BFD session establishment modes

BFD Session Establishment Mode	Description
Static mode	BFD session parameters, such as the local and remote discriminators, are manually configured and delivered for BFD session establishment. **NOTE** In static mode, configure unique local and remote discriminators for each BFD session. This mode prevents incorrect discriminators from affecting BFD sessions that have correct discriminators and prevents BFD sessions from alternating between Up and Down.
Dynamic mode	When a BFD session is dynamically established, the system processes the local and remote discriminators as follows: • Dynamically allocates the local discriminator. When a system triggers the dynamic establishment of a BFD session, the system allocates a dynamic discriminator as the local discriminator of the BFD session. Then, the system sends a BFD control packet with Your Discriminator set to 0 to the peer for session negotiation. • Automatically learns the remote discriminator. The local end of a BFD session sends a BFD control packet with Your Discriminator set to 0 to the remote end. After the remote end receives the packet, it checks whether the value of Your Discriminator in this packet is the same as the value of its My Discriminator. If the value of Your Discriminator matches that of My Discriminator, the remote end learns the value of My Discriminator of the local end and obtains its Your Discriminator.

BFD Session Management

A BFD session has the following states:

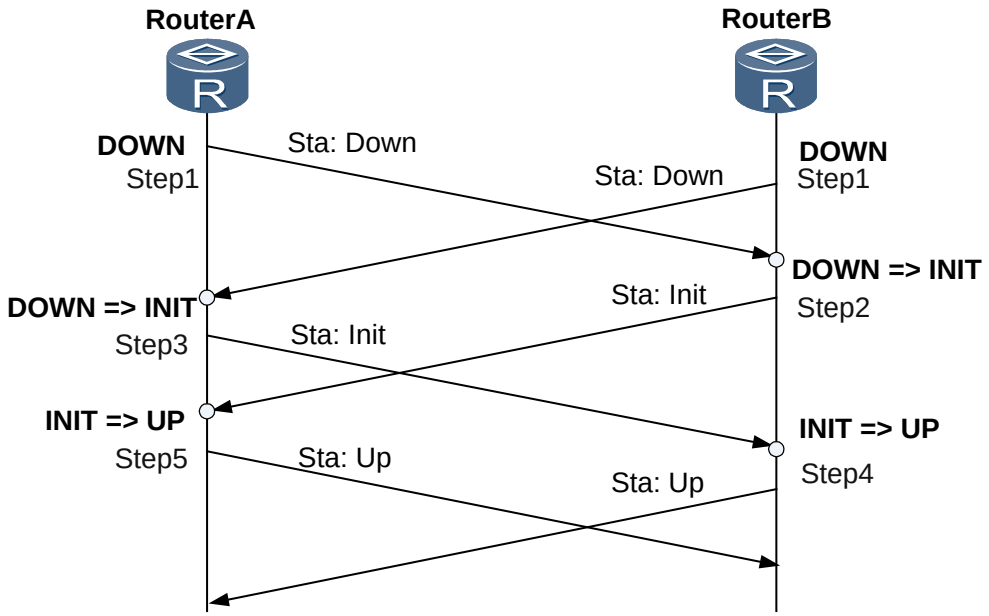
- Down: indicates that the BFD session is in the Down state or has just been set up.
- Init: indicates that the local system can communicate with the remote system, and the local system expects a BFD session to go Up.
- Up: indicates that the BFD session is set up successfully.
- AdminDown: indicates that the BFD session is in the administratively Down state.

The session status is conveyed in the State field of a BFD control packet. The system changes the session status based on the local session status and the received session status of the peer.

When a BFD session is to be set up or deleted, the BFD state machine implements a three-way handshake to ensure that both two systems are aware of the status change.

Figure 19-61 shows the state transition process in establishment of a BFD session.

Figure 19-61 BFD session state transition



1. BFD configured on both Router A and Router B independently starts state machines. The initial status of BFD state machines is Down. Router A and Router B send BFD control packets with the State field set to Down. If BFD sessions are established in static mode, the value of Your Discriminator in BFD control packets is manually specified. If BFD sessions are established in dynamic mode, the value of Your Discriminator is set to 0.
2. After receiving a BFD control packet with the State field set to Down, Router B switches the session status to Init and sends a BFD control packet with the State field set to Init.

NOTE

After the local BFD session status of Router B changes to Init, Router B no longer processes the received BFD control packets with the State field set to Down.

3. The BFD session status change of Router A is the same as that of Router B.

4. After receiving a BFD control packet with the State field set to Init, Router B changes the local session status to Up.
5. The BFD session status change of Router A is the same as that of Router B.

19.10.3 Application Environment

BFD for IP

A BFD session is established on an IP link to fast detect faults.

BFD can detect single-hop and multi-hop IP links.

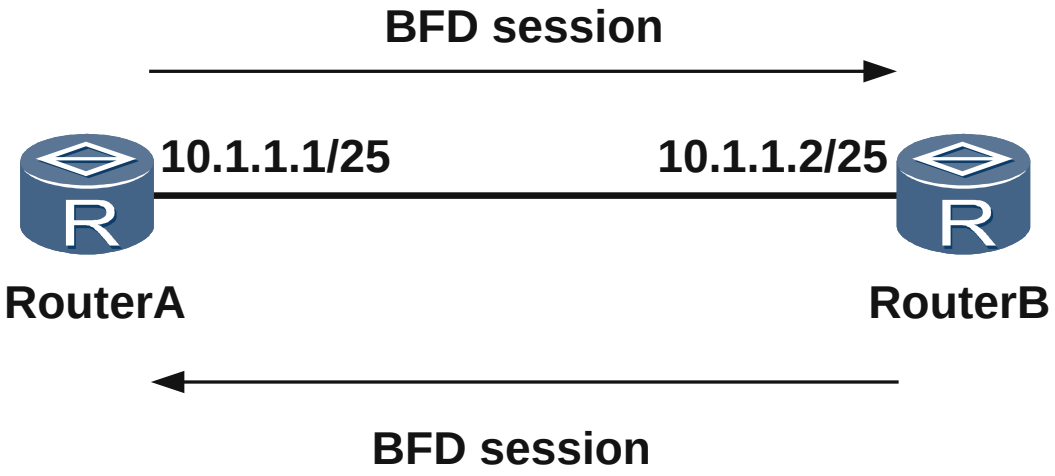
- Single-hop BFD detects IP route connectivity between directly-connected systems. The single hop refers to an IP hop. Between these two systems, only one BFD session can be set up for a specified data protocol on an interface.
- Multi-hop BFD detects any paths between systems. A path may span multiple hops or may partially overlap.

BFD for IP Applications

Example 1

Figure 19-62 shows a single-hop BFD session detecting a path between directly-connected devices. The BFD session is bound to the outgoing interface.

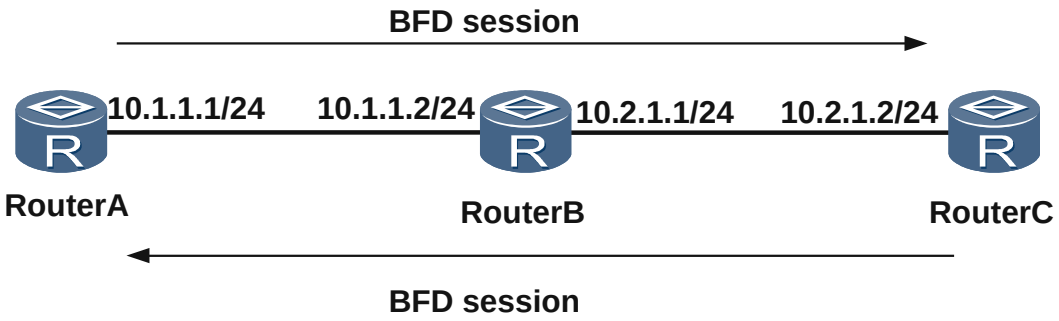
Figure 19-62 Single-hop BFD for IP



Example 2

Figure 19-63 shows a multi-hop BFD session detecting a path between Router A and Router C. The BFD session is bound to the peer IP address but not the outgoing interface.

Figure 19-63 Multi-hop BFD for IP



BFD for USR

BFD for Unicast Static Route (USR) is used to detect IPv4 USRs. After a BFD session is bound to an IPv4 USR, link failures can be detected more quickly.

Unlike dynamic routing protocols, USRs do not have a detection mechanism. If a fault occurs on a network, an administrator needs to handle it manually. In BFD for USR, BFD sessions are bound to IPv4 USRs in a public network and are used to detect the link status of the IPv4 USR.

Each BFD session is bound to a single IPv4 USR. When a BFD session detects a fault (for example, the link changes from Up to Down) on a link of the USR, BFD reports the fault to the routing management module. Then, the RM sets the USR as "inactive" (indicating that the route is unavailable and is deleted from the IP routing table).

When the BFD session bound to the USR is successfully set up or the link of the USR recovers from the fault (that is, the link changes from Down to Up), BFD reports the event to the RM and the RM sets the USR as "active" (indicating that the route is available and has been added to the IP routing table).

BFD for OSPF

A link fault or change in topology may lead to rerouting in a network. Quick convergence of a routing protocol is important for improving network availability. A feasible solution is to fast detect the fault and immediately notify the routing protocol of the fault.

In BFD for OSPF, OSPF is associated with a BFD session. The BFD session fast detects a link fault and notifies OSPF of the fault. In this manner, OSPF speeds up responses to changes in network topology.

Table 19-12 shows convergence speed statistics when OSPF is and is not associated with a BFD session.

Table 19-12 OSPF convergence speed statistics

Associated with BFD	Link Fault Detection Mechanism	Convergence Speed
No	OSPF Hello keepalive timer timeout	Within seconds
Yes	BFD session in the Down state	Within milliseconds

Figure 19-64 BFD for OSPF networking diagram

As shown in Figure 19-64, Router A sets up OSPF neighbor relationships with Router C and Router D. The outbound interface VLANIF 10 on Router A is connected to Router B through Router C. When the neighbor state is Full, BFD is notified of the status and starts to set up a BFD session.

1. When a fault occurs on the link between Router A and Router C, the BFD session detects the fault and notifies Router A.
2. Router A processes the neighbor-Down event and recalculates routes. Then, the outbound interface changes to VLANIF 20 on Router A, which is connected Router B through Router D.

BFD for OSPFv3

Definition

Bidirectional Forwarding Detection (BFD) is a mechanism used to detect faults of communications between forwarding engines.

To be specific, BFD detects connectivity of a data protocol on a path between two systems. The path can be a physical link, a logical link, or a tunnel.

BFD for OSPFv3 associates BFD with OSPFv3. BFD fast detects a link fault and then notifies OSPFv3 of the fault. This speeds up OSPFv3's response to the change of the network topology.

Purpose

A link fault or the topology change causes Routers to recalculate routes. Therefore, the convergence of routing protocols must be as quick as possible to improve network performance.

Link faults are inevitable. Therefore, fast detecting faults and notifying routing protocols of the faults is a feasible solution to immediately rectify link faults. After BFD is associated with routing protocols, BFD can speed up the convergence of routing protocols if a link fault occurs.

Principles

Figure 19-65 BFD for OSPFv3

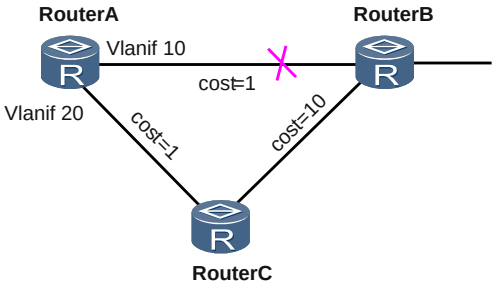


Figure 19-65 shows the principle of BFD for OSPFv3.

1. OSPFv3 neighbor relationships are established between these three Routers.
2. After a neighbor relationship becomes Full, this triggers BFD to establish a BFD session.
3. The outbound interface on Router A connected to Router B is Vlanif10. If the link fails, BFD detects the fault and then notifies Router A of the fault.
4. Router A processes the event that a neighbor relationship becomes Down and re-calculates routes. After calculation, the outbound interface is Vlanif20 passes through Router C and then reaches Router B.

BFD for IS-IS

Generally, the interval at which the Intermediate System to Intermediate System (IS-IS) protocol sends Hello messages is 10 seconds. If a device does not receive any Hello message from its neighbor within three Hello intervals, the device deletes the neighbor. Therefore, it takes a device a number of seconds to detect that a neighbor is Down. This leads to the loss of a large number of packets in a high-speed network.

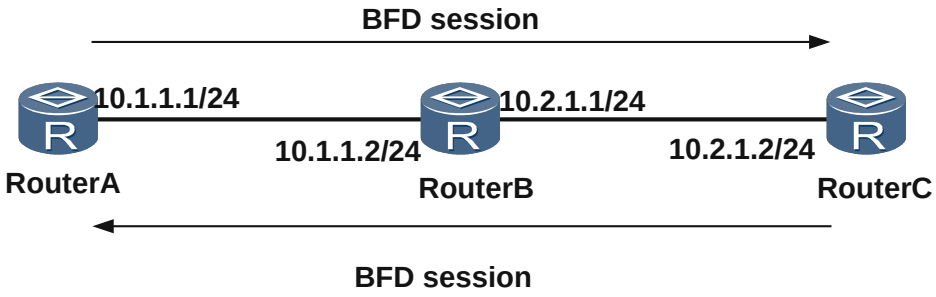
In BFD for IS-IS, the establishment of a BFD session is dynamically triggered by IS-IS but not configured manually. When detecting a fault, the BFD session notifies IS-IS of the fault through the Routing Management Module (RM). IS-IS processes the neighbor-Down event and quickly sends the link state PDU (LSP), and performs the partial route calculation (PRC). In this manner, IS-IS routes fast converge.

The BFD fault detection interval is at the millisecond level. Instead of replacing the IS-IS Hello mechanism, BFD works with IS-IS to detect the adjacency fault more quickly. In addition, BFD instructs IS-IS to recalculate routes, ensuring correct packet forwarding.

The RM allows IS-IS and BFD to interact with each other. Through the RM, IS-IS instructs BFD to dynamically set up or delete BFD sessions. The BFD event messages are also delivered to IS-IS through the RM.

BFD for IS-IS Applications

Figure 19-66 BFD for IS-IS networking diagram



After BFD is enabled on RouterA, RouterB, and RouterC, the BFD session can quickly detect faults on the link between RouterA and RouterB, and notify IS-IS through the RM. Then, IS-IS sets the neighbor status to Down to trigger the IS-IS topology calculation. In addition, IS-IS updates LSPs to ensure that RouterC (RouterB's neighbor) can receive the updated LSPs from RouterB in time. This implements fast network topology convergence.

BFD for BGP

The Border Gateway Protocol (BGP) periodically sends Keepalive messages to its peer to monitor the neighbor status. This detection process lasts more than 1 second. When the data is transmitted at gigabit rates, a large amount of data will be discarded, which cannot meet the requirement for carrier-class reliability.

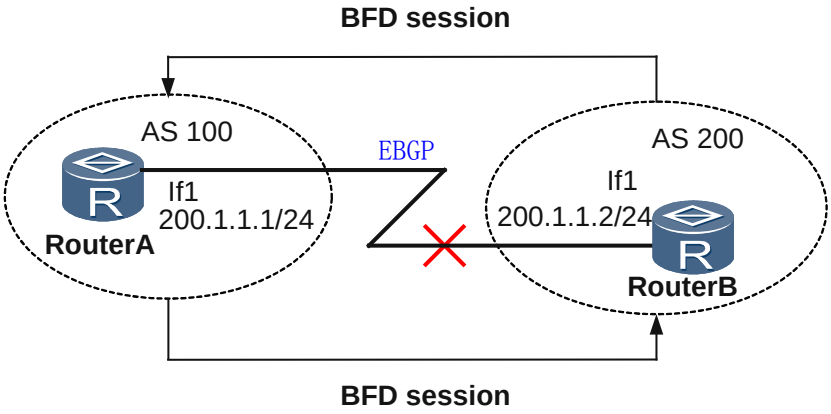
BFD for BGP was developed to compensate for this shortcoming. The BFD session can fast detect a fault on a link between BGP peers and notify BGP, ensuring fast convergence.

NOTE

By default, a multi-hop BGP session is established between Huawei devices that set up an IBGP peer relationship. A BFD for IGP session and A BFD for IBGP session cannot be both set up between a Huawei device and a non-Huawei device that sets up a single-hop BGP session with its peer by default. In such a situation, setting up only A BFD for IGP session or A BFD for IBGP session between the Huawei and non-Huawei devices is recommended.

BFD for BGP Applications

Figure 19-67 BFD for BGP networking



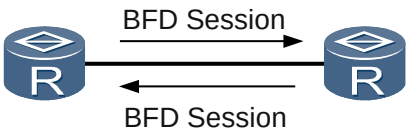
As shown in Figure 19-67, RouterA belongs to AS 100 and RouterB belongs to AS 200. RouterA and RouterB are directly connected through the External Border Gateway Protocol (EGBP). A BFD session is established to detect the BGP neighbor relationship between RouterA and RouterB. When the link between RouterA and RouterB is faulty, the BFD session can quickly detect the fault and notify BGP.

BFD for RSVP

BFD monitors RSVP neighbor relationships. When a Layer 2 device (a hub for example) exists between RSVP neighboring nodes, the two nodes can detect a link fault only using the Hello mechanism in seconds. This process results in the loss of lots of data. BFD for RSVP rapidly detects faults in a link between RSVP neighboring nodes within milliseconds. BFD for RSVP applies to TE FRR networks, on which Layer 2 devices exist on a primary CR-LSP between the PLR and its RSVP neighboring node.

BFD for RSVP is an IP-layer detection. Only a single-hop BFD session can be set up between RSVP neighboring nodes.

Figure 19-68 BFD for RSVP



A BFD session for RSVP is set up to monitor the link between RSVP neighbors. The RSVP module can rapidly detect a link failure.

BFD for RSVP can share BFD sessions with BFD for Open Shortest Path First (OSPF), BFD for Intermediate System to Intermediate System (IS-IS), or BFD for Border Gateway Protocol (BGP). The local node selects the smallest values of parameters between the two ends of the shared BFD session as local BFD parameters. The parameters include the interval at which BFD packets are sent, interval at which BFD packets are received, and local detection multiplier.

19.10.4 Configuring the BFD

This topic describes how to configure the BFD on the MA5600T/MA5603T/MA5608T.

Context

Bidirectional Forwarding Detection (BFD) protocol is a draft standardized by the Internet Engineering Task Force (IETF). BFD rapidly detects faults and monitors the forwarding and connectivity of links or IP routes of the network by quickly sending BFD control packets (the UDP packets in a specified format) at intervals between two nodes.

BFD provides the following functions:

- Allows fault detection with light load and high speed for paths between the neighboring forwarding engines.
- Provides a single mechanism to detect any medium and protocol layer in real time.

Configuring BFD Sessions

A bidirectional forwarding detection (BFD) session rapidly detects faults in links over a network.

Context

In the BFD detection mechanism, two systems set up a BFD session, and periodically send BFD control packets along the path between them. If one system does not receive BFD control packets within a specified period, the system considers that a fault occurs on the path.

BFD uses the local and remote discriminators to differentiate multiple BFD sessions between the same pair of systems. Based on the differences in methods of creating the local and the remote discriminators, MA5600T/MA5603T/MA5608T supports the following types of BFD sessions:

- Static BFD sessions with manually-specified discriminators

  The local and remote discriminators must be set manually. The discriminators on the remote end must also be manually specified.
- Static BFD sessions with automatically-negotiated discriminators

  If a dynamic BFD session is used by a remote device, a static BFD session with automatically negotiated discriminators must be created on a local device to interwork with the remote device and support the BFD for static routes. The discriminators on the remote end can be automatically negotiated or a dynamic BFD session can be established on the remote end.
- BFD sessions dynamically triggered by protocols, where no local or remote discriminator needs to be set:
  - BFD sessions with dynamically-allocated local discriminators.
  - BFD sessions with self-learned remote discriminators.

Procedure

Enable BFD globally.

1. Run the **bfd** command to enable BFD globally and enter the BFD mode.

BFD must be enabled globally before configurations relevant to BFD are performed. By default, BFD is disabled globally.

2. Run the **quit** command to quit the BFD mode.

Step 1 Create a BFD session. Select one of the following steps depending on the type of link to be checked by BFD.

- For an IPv4 link

  Run the **bfd bind peer-ip** command to create a BFD session.

  If the **bfd bind peer-ip source-ip auto** command is run, a BFD session is set up through automatic negotiations over discriminators. The device on which such a BFD session is created can interoperate with another device on which a dynamic BFD is set up. This command is mainly used to configure BFD sessions for IPv4 static routes.

- For an IPv6 link

  Run the **bfd bind peer-ipv6** command to create a BFD session.

  If the **bfd bind peer-ipv6 source-ipv6 auto** command is run, a BFD session is set up through automatic negotiations over discriminators. The device on which such a BFD session is created can interoperate with another device on which a dynamic BFD is set up. This command is mainly used to configure BFD sessions for IPv6 static routes.

Pay attention to the following points:

- If a single-hop BFD session is to be set up on an interface for the first time, the interface and its peer address must be bound to the BFD session. The bindings cannot be modified after the BFD session is successfully created.
- If a multi-hop BFD session is to be set up on an interface for the first time, the peer address must be bound to the BFD session. The bindings cannot be modified after the BFD session is successfully created.
- During BFD configuration items are being created, the system checks only the format, not the correctness, of an IP address. Either an incorrect peer or source IP address leads to a failure in creating a BFD session.

Step 2 Configure the discriminators.

1. Run the **discriminator local** *discr-value* command to configure a local discriminator.
2. Run the **discriminator remote** *discr-value* command to configure a remote discriminator.

The local discriminator set on a device is equal to the remote discriminator set on a remote device, and the remote discriminator set on the local device is equal to the local discriminator set on the remote device. If the discriminators on the device and the remote device do not match, the session cannot be created. After the local and remote discriminators are set, they cannot be changed.

Step 3 (Optional) Configure the BFD parameters.

Select the following desired operations:

- Modify the detection time.
  - Run the **min-tx-interval** command to configure the interval for sending BFD packets.
  - Run the **min-rx-interval** command to configure the interval for receiving BFD packets.
- Run the **detect-multiplier** command to configure the local detection multiplier.

- Run the **description** command to add the description of a BFD session. Descriptions of BFD sessions help you distinguish between various BFD sessions.

  The **description** command takes effect only on the statically configured BFD sessions, rather than the BFD sessions that are dynamically configured or the BFD sessions that are set up through automatic negotiations over discriminators.
- Run the **tos-exp** command to configure the priority of the BFD packet. By default, the highest priority 7 is adopted. When the system is congested, the BFD packet with higher priority can be sent first.

  You can configure the priority in static BFD mode but not in dynamic BFD mode.
- Run the **wtr** command to configure the time of waiting for recovery of the BFD session. By default, the value is 0, indicating no waiting.

  The BFD session is unidirectional. The detection is performed by BFD parameters configured on both ends respectively. If wait-to-recovery (WTR) is needed, configure it on two ends manually. Or, when the status of the session on one end changes, the applications on both ends can find that the states of the BFD sessions are inconsistent.

Step 4 Run the **commit** command to commit the configuration.

After necessary parameters, such as local and remote discriminators, are configured for a BFD session, the **commit** command must be run to make the configuration take effect.

After a BFD session has been created, to modify a parameter, run a corresponding command (such as **min-tx-interval**, **min-rx-interval**, **detect-multiplier**, **description**, **tos-exp**, or **wtr**). The modification takes effect immediately without the **commit** command configured.

Step 5 Query the BFD session information and BFD session statistics.

- Run the **display bfd configuration** command to query the BFD configuration.
- Run the **display bfd interface** command to query the BFD configuration on an interface.
- Run the **display bfd session** command to query the BFD session information.
- Run the **display bfd statistics** command to query the BFD global statistics.
- Run the **display bfd statistics session** command to query the BFD session statistics.

----End

Example

Assume that the peer IP address is 10.1.1.1/24, BFD session name is test, the local discriminator is 100, the remote discriminator is 200, the minimum transmit interval and minimum receive interval of BFD control packets are both 10 milliseconds, the local detection multiplier is 3 (default value), VLAN 10 is created, and the IP address of VLAN interface 10 is configured. To configure BFD single-hop detection on VLAN interface 10, run the following commands:

```
huawei(config)#bfd
huawei(config-bfd)#quit
huawei(config)#bfd test bind peer-ip 10.1.1.1 interface vlanif 10
huawei(config-config-bfd-session-test)#discriminator local 100
huawei(config-config-bfd-session-test)#discriminator remote 200
huawei(config-config-bfd-session-test)#min-tx-interval 10
huawei(config-config-bfd-session-test)#min-rx-interval 10
huawei(config-config-bfd-session-test)#commit
```

Configuring BFD for Static Routes

The MA5600T/MA5603T/MA5608T supports detecting the fault of a static route by using the BFD. This topic describes how to configure the BFD link detection based on an example network.

Prerequisites

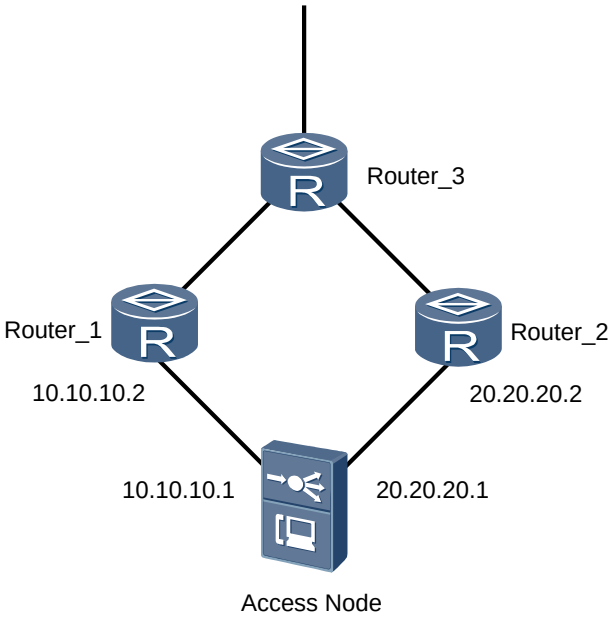
The BFD function must be enabled globally on the MA5600T/MA5603T/MA5608T.

Networking

Figure 19-69 shows an example network of the BFD for Static Routes.

Different static routes exist between the MA5600T/MA5603T/MA5608T and Router_3 through Router_1 and Router_2, and the BFD session is bound to the static route. When one link is faulty, the BFD session notifies the bound route for route switching.

Figure 19-69 Example network of the BFD for Static Routes



Data Plan

Table 19-13 provides the data plan for configuring the BFD for Static Routes.

Table 19-13 Data plan for configuring the BFD for Static Routes

Item	Data	Remarks
MA5600T/MA5603T/MA5608T	Upstream ports: 0/19/0 and 0/19/1	-
VLAN	VLAN ID: 30 VLAN type: Smart VLAN	-

Item	Data	Remarks
	IP address of the Layer 3 interface: 10.10.10.1/24	
	VLAN ID: 40 VLAN type: Smart VLAN IP address of the Layer 3 interface: 20.20.20.1/24	-
BFD session	Session name: ToRouter_1 IP address of the peer interface: 10.10.10.2/24 Minimum transmit interval: 10 ms Minimum receive interval: 10 ms Detection multiplier: 3 Identifier: auto-negotiation	-
	Session name: ToRouter_2 IP address of the peer interface: 20.20.20.2/24 Minimum transmit interval: 10 ms Minimum receive interval: 10 ms Detection multiplier: 3 Identifier: auto-negotiation	-
Static route	Destination address: 30.30.30.1/24 Priority of the static route with next hop Router_1: 2 Priority of the static route with next hop Router_2: 6	-
Requirements for the upper-layer device	Router_1: • IP address of the Layer 3 interface: see the example network • VLAN ID: 30 • BFD session parameters: consistent with the parameters of the MA5600T/MA5603T/MA5608T	For details about the configuration of the routers, see the corresponding configuration guide.
	Router_2: • IP address of the Layer 3 interface: see the example network • VLAN ID: 40 • BFD session parameters: consistent with the parameters of the MA5600T/MA5603T/MA5608T	

Procedure

Create VLANs and add upstream ports to the VLANs.

```
huawei(config)#vlan 30 smart
huawei(config)#port vlan 30 0/19 0
```

```
huawei(config)#vlan 40 smart
huawei(config)#port vlan 40 0/19 1
```

Step 1 Configure the IP address of the Layer 3 interface of the VLAN.

```
huawei(config)#interface vlanif 30
huawei(config-if-vlanif30)#ip address 10.10.10.1 24
huawei(config-if-vlanif30)#quit
huawei(config)#interface vlanif 40
huawei(config-if-vlanif40)#ip address 20.20.20.1 24
huawei(config-if-vlanif40)#quit
```

Step 2 Configure the BFD sessions.

You can configure BFD sessions only after the global BFD function is enabled.

```
huawei(config)#bfd
huawei(config-bfd)#quit
huawei(config)#bfd ToRouter_1 bind peer-ip 10.10.10.2 source-ip 10.10.10.1 auto
huawei(config-bfd-session-torouter_1)#min-rx-interval 10
huawei(config-bfd-session-torouter_1)#min-tx-interval 10
huawei(config-bfd-session-torouter_1)#detect-multiplier 3
huawei(config-bfd-session-torouter_1)#commit
huawei(config-bfd-session-torouter_1)#quit
huawei(config)#bfd ToRouter_2 bind peer-ip 20.20.20.2 source-ip 20.20.20.1 auto
huawei(config-bfd-session-torouter_2)#min-rx-interval 10
huawei(config-bfd-session-torouter_2)#min-tx-interval 10
huawei(config-bfd-session-torouter_2)#detect-multiplier 3
huawei(config-bfd-session-torouter_2)#commit
huawei(config-bfd-session-torouter_2)#quit
```

Step 3 Bind the BFD sessions to the static routes.

```
huawei(config)#ip route-static 30.30.30.1 24 10.10.10.2 preference 2 track bfd-session
ToRouter_1
huawei(config)#ip route-static 30.30.30.1 24 20.20.20.2 preference 6 track bfd-session
ToRouter_2
```

Step 4 Save the data.

```
huawei(config)#save
```

----End

Result

BFD sessions ToRouter_1 and ToRouter_2 are in the up state. The priority of the route to which ToRouter_1 is bound takes effect and carries services because it has a higher priority. When a faulty link is detected, BFD session ToRouter_1 turns to the down state, which triggers the deactivation of the bound route. In this case, the route to which ToRouter_2 is bound takes effect and carries services.

Configuring BFD for OSPF

The MA5600T/MA5603T/MA5608T can detect the fault of a dynamic route by using the bidirectional forwarding detection (BFD). This topic describes how to configure the BFD link detection based on the dynamic routing protocol open shortest path first (OSPF).

Prerequisites

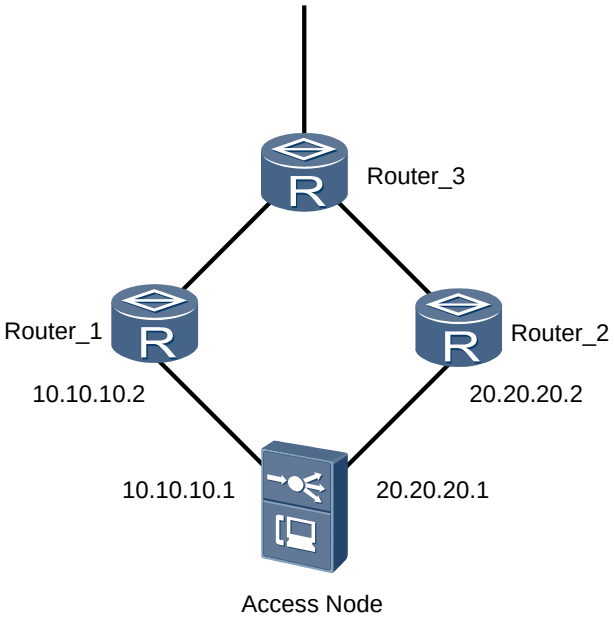
The BFD function must be globally enabled on the MA5600T/MA5603T/MA5608T.

Networking

Figure 19-70 shows an example network of the BFD for OSPF.

Dynamic routes between the MA5600T/MA5603T/MA5608T and Router_1, Router_2 are generated through OSPF. The BFD session is bound to the OSPF route. When one link is faulty, the BFD session reports that the bound OSPF neighbor is down, switching the route.

Figure 19-70 Example network of the BFD for OSPF



Data Plan

Table 19-14 provides the data plan for configuring the BFD for OSPF.

Table 19-14 Data plan for configuring the BFD for OSPF

Item	Data	Remarks
MA5600T/MA5603T/MA5608T	Upstream ports: 0/19/0 and 0/19/1	-
VLAN	Virtual local area network (VLAN) ID: 30 VLAN type: Smart VLAN IP address of the Layer 3 interface: 10.10.10.1/24	-
	VLAN ID: 40 VLAN type: Smart VLAN IP address of the Layer 3 interface: 20.20.20.1/24	-

Item	Data	Remarks
BFD session	Minimum transmit interval: 10 ms Minimum receive interval: 10 ms Detection multiplier: 3	-
Requirements for the upper-layer device	Router_1: • IP address of the Layer 3 interface: see the example network • VLAN ID: 30 • OSPF: enabled • BFD session parameters: consistent with the parameters of the MA5600T/MA5603T/MA5608T	For details about the configuration of the router, see the corresponding configuration guide.
	Router_2: • IP address of the Layer 3 interface: see the example network • VLAN ID: 40 • OSPF: enabled • BFD session parameters: consistent with the parameters of the MA5600T/MA5603T/MA5608T	

Procedure

Create VLANs and add upstream ports to the VLANs.

```
huawei(config)#vlan 30 smart
huawei(config)#port vlan 30 0/19 0
huawei(config)#vlan 40 smart
huawei(config)#port vlan 40 0/19 1
```

Step 1 Configure the IP address of the Layer 3 interface of the VLAN.

```
huawei(config)#interface vlanif 30
huawei(config-if-vlanif30)#ip address 10.10.10.1 24
huawei(config-if-vlanif30)#quit
huawei(config)#interface vlanif 40
huawei(config-if-vlanif40)#ip address 20.20.20.1 24
huawei(config-if-vlanif40)#quit
```

Step 2 Configure basic OSPF functions.

```
huawei(config)#ospf 1
huawei(config-ospf-1)#area 0
huawei(config-ospf-1-area-0.0.0.0)#network 10.10.10.0 0.0.0.255
huawei(config-ospf-1-area-0.0.0.0)#network 20.20.20.0 0.0.0.255
huawei(config-ospf-1-area-0.0.0.0)#quit
huawei(config-ospf-1)#quit
```

Step 3 Configure BFD function on the OSPF interface.

You can configure BFD sessions only after the global BFD function is enabled.

```
huawei(config)#bfd
huawei(config-bfd)#quit
huawei(config)#interface vlanif 30
huawei(config-if-vlanif30)#ospf bfd enable
huawei(config-if-vlanif30)#ospf bfd min-rx-interval 10 min-tx-interval 10
detect-multiplier 3
huawei(config-if-vlanif30)#ospf cost 30
huawei(config-if-vlanif30)#quit
huawei(config)#interface vlanif 40
huawei(config-if-vlanif40)#ospf bfd enable
huawei(config-if-vlanif40)#ospf bfd min-rx-interval 10 min-tx-interval 10
detect-multiplier 3
huawei(config-if-vlanif30)#ospf cost 40
huawei(config-if-vlanif40)#quit
```

Step 4 Save the data.

```
huawei(config)#save
```

----End

Result

After establishing the neighbor relation with each router through OSPF, the MA5600T/MA5603T/MA5608T automatically creates two BFD sessions. When the active link is faulty, its bound BFD session is down, which triggers the OSPF neighbor relation to be down. Therefore, the route is switched to the standby link.

Run the **display ospf bfd session** command to query the BFD session information.

Configuring BFD6 for OSPFv3

The MA5600T/MA5603T/MA5608T can detect the fault of a dynamic route by using the bidirectional forwarding detection (BFD). This topic describes how to configure the BFD link detection based on the dynamic routing protocol open shortest path first (OSPFv3).

Prerequisites

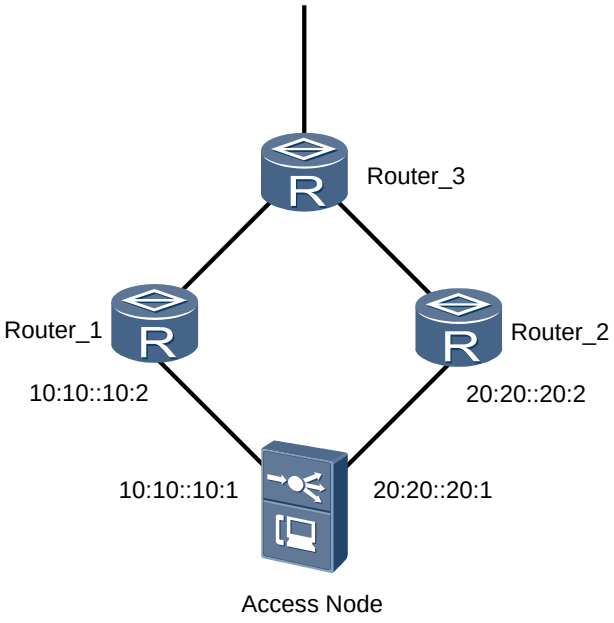
The BFD function must be globally enabled on the MA5600T/MA5603T/MA5608T.

Networking

Figure 19-71 shows an example network of the BFD6 for OSPFv3.

Dynamic routes between the MA5600T/MA5603T/MA5608T and Router_1, Router_2 are generated through OSPFv3. The BFD session is bound to the OSPFv3 route. When one link is faulty, the BFD session reports that the bound OSPFv3 neighbor is down, switching the route.

Figure 19-71 Example network of the BFD6 for OSPFv3



Data Plan

Table 19-15 provides the data plan for configuring the BFD6 for OSPFv3.

Table 19-15 Data plan for configuring the BFD6 for OSPFv3

Item	Data	Remarks
MA5600T/MA5603T/MA5608T	Upstream ports: 0/19/0 and 0/19/1	-
VLAN	Virtual local area network (VLAN) ID: 30 VLAN type: Smart VLAN IPv6 address of the Layer 3 interface: 10:10::10:1/64	-
	VLAN ID: 40 VLAN type: Smart VLAN IPv6 address of the Layer 3 interface: 20:20::20:1/64	-
BFD session	Minimum transmit interval: 10 ms Minimum receive interval: 10 ms Detection multiplier: 3	-
Requirements for the upper-layer device	Router_1: • IPv6 address of the Layer 3 interface: see the example network • VLAN ID: 30 • OSPFv3: enabled • BFD session parameters: consistent with the	For details about the configuration of the router, see the correspondin g

Item	Data	Remarks
	parameters of the MA5600T/MA5603T/MA5608T	configuration guide.
	Router_2: • IPv6 address of the Layer 3 interface: see the example network • VLAN ID: 40 • OSPFv3: enabled • BFD session parameters: consistent with the parameters of the MA5600T/MA5603T/MA5608T	

Procedure

Create VLANs and add upstream ports to the VLANs.

```
huawei(config)#vlan 30 smart
huawei(config)#port vlan 30 0/19 0
huawei(config)#vlan 40 smart
huawei(config)#port vlan 40 0/19 1
```

Step 1 Configure the IPv6 address of the Layer 3 interface of the VLAN.

```
huawei(config)#ipv6
huawei(config)#interface vlanif 30
huawei(config-if-vlanif30)#ipv6 enable
huawei(config-if-vlanif30)#ipv6 address 10:10::10:1 64
huawei(config-if-vlanif30)#quit
huawei(config)#interface vlanif 40
huawei(config-if-vlanif40)#ipv6 enable
huawei(config-if-vlanif40)#ipv6 address 20:20::20:1 64
huawei(config-if-vlanif40)#quit
```

Step 2 Configure basic OSPFv3 functions.

```
huawei(config)#ospfv3
huawei(config-ospfv3-1)#quit
huawei(config)#interface vlanif 30
huawei(config-if-vlanif30)#ospfv3 1 area 0
huawei(config-if-vlanif30)#quit
huawei(config)#interface vlanif 40
huawei(config-if-vlanif40)#ospfv3 1 area 0
huawei(config-if-vlanif40)#quit
```

Step 3 Configure BFD function on the OSPFv3 interface.

You can configure BFD sessions only after the global BFD function is enabled.

```
huawei(config)#bfd
huawei(config-bfd)#quit
huawei(config)#interface vlanif 30
huawei(config-if-vlanif30)#ospfv3 bfd enable
```

```
huawei(config-if-vlanif30)#ospfv3 bfd min-rx-interval 10 min-tx-interval 10
detect-multiplier 3
huawei(config-if-vlanif30)#ospfv3 cost 30
huawei(config-if-vlanif30)#quit
huawei(config)#interface vlanif 40
huawei(config-if-vlanif40)#ospfv3 bfd enable
huawei(config-if-vlanif40)#ospfv3 bfd min-rx-interval 10 min-tx-interval 10
detect-multiplier 3
huawei(config-if-vlanif40)#ospfv3 cost 40
huawei(config-if-vlanif40)#quit
```

Step 4 Save the data.

```
huawei(config)#save
```

----End

Result

After establishing the neighbor relation with each router through OSPFv3, the MA5600T/MA5603T/MA5608T automatically creates two BFD sessions. When the active link is faulty, its bound BFD session is down, which triggers the OSPFv3 neighbor relation to be down. Therefore, the route is switched to the standby link.

Run the **display ospfv3 bfd session** command to query the BFD session information.

Configuring BFD for IS-IS

The access device can detect the fault of a dynamic route by using the bidirectional forwarding detection (BFD). This topic describes how to configure the BFD link detection based on the dynamic routing protocol intermediate system to intermediate system (IS-IS).

Context

To accelerate IS-IS convergence speed when the link status changes, you can configure BFD on the IS-IS link. The access device supports configuration of static and dynamic BFD for IS-IS. When BFD sessions are configured in both methods, the static BFD session takes precedence over the dynamic BFD session.

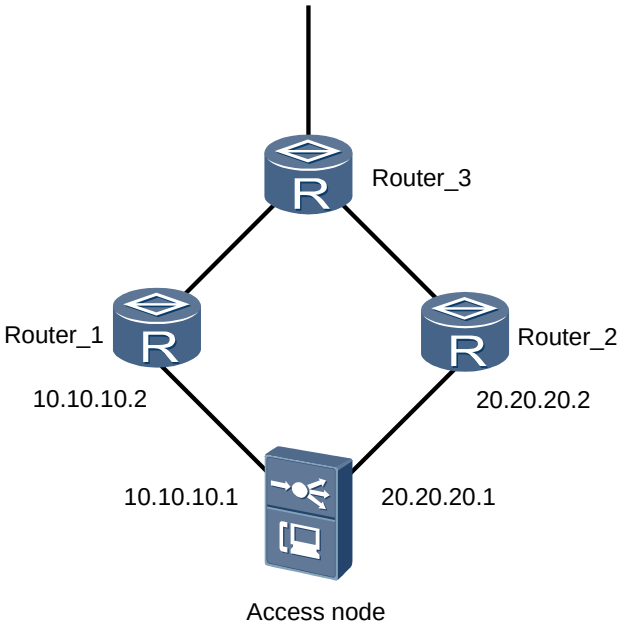
- Static BFD refers to configuring BFD session parameters manually including local and remote identifiers and delivering BFD session setup requests manually.
- Dynamic BFD refers to that routing protocols dynamically trigger the establishment of BFD sessions. When setting up new neighbor relationship, routing protocols send parameters of neighbors and detection parameters (including source and destination IP addresses) to the BFD module. BFD then sets up sessions according to the received parameters between neighbors. Dynamic BFD is more flexible than static BFD.

Networking

Figure 19-72 shows an example network of the BFD for IS-IS.

Dynamic routes between the access node and Router_1, Router_2 are generated through IS-IS. The BFD session is bound to the IS-IS route. When one link is faulty, the BFD session reports that the bound IS-IS neighbor is down, switching the route.

Figure 19-72 Example network of the BFD for IS-IS



Data Plan

Table 19-16 provides the data plan for configuring the BFD for IS-IS.

Table 19-16 Data plan for configuring the BFD for IS-IS

Item	Data	Remarks
Access node	Upstream ports: 0/19/0 and 0/19/1	-
VLAN	Virtual local area network (VLAN) ID: 30 VLAN type: Smart VLAN IP address of the Layer 3 interface: 10.10.10.1/24	-
	VLAN ID: 40 VLAN type: Smart VLAN IP address of the Layer 3 interface: 20.20.20.1/24	-
BFD session	Static BFD session name: ToRouter_1 IP address of the peer interface: 10.10.10.2/24 Local discriminator: 1 Remote discriminator: 2 Minimum transmit interval: 10 ms Minimum receive interval: 10 ms Detection multiplier: 3	-
	Static BFD session name: ToRouter_2 IP address of the peer interface: 20.20.20.2/24	

Item	Data	Remarks
	Local discriminator: 3 Remote discriminator: 4 Minimum transmit interval: 10 ms Minimum receive interval: 10 ms Detection multiplier: 3	
IS-IS	Network Entity Title (NET): aa.1111.1111.1111.00 Link cost of VLAN interface 30: 30 Link cost of VLAN interface 40: 40	-
Requirements for the upper-layer device	Router_1: • IP address of the Layer 3 interface: see the example network • VLAN ID: 30 • IS-IS: enabled • Local discriminator of the static BFD session: 2 • Remote discriminator of the static BFD session: 1 • Dynamic BFD session parameters: consistent with the parameters of the access node	For details about the configuration of the router, see the corresponding configuration guide.
	Router_2: • IP address of the Layer 3 interface: see the example network • VLAN ID: 40 • IS-IS: enabled • Local discriminator of the static BFD session: 4 • Remote discriminator of the static BFD session: 3 • Dynamic BFD session parameters: consistent with the parameters of the access node	

Procedure

Create VLANs and add upstream ports to the VLANs.

```
huawei(config)#vlan 30 smart
huawei(config)#port vlan 30 0/19 0
huawei(config)#vlan 40 smart
huawei(config)#port vlan 40 0/19 1
```

Step 1 Configure the IP address of the Layer 3 interface of the VLAN.

```
huawei(config)#interface vlanif 30
huawei(config-if-vlanif30)#ip address 10.10.10.1 24
huawei(config-if-vlanif30)#quit
huawei(config)#interface vlanif 40
huawei(config-if-vlanif40)#ip address 20.20.20.1 24
huawei(config-if-vlanif40)#quit
```

Step 2 Configure IS-IS.

```
huawei(config)#isis
huawei(config-isis-1)#network-entity aa.1111.1111.1111.00
huawei(config-isis-1)#quit
huawei(config)#interface vlanif 30
huawei(config-if-vlanif30)#isis enable
huawei(config-if-vlanif30)#isis cost 30
huawei(config-if-vlanif30)#quit
huawei(config)#interface vlanif 40
huawei(config-if-vlanif40)#isis enable
huawei(config-if-vlanif40)#isis cost 40
huawei(config-if-vlanif40)#quit
```

Step 3 Configure static BFD for IS-IS.

```
huawei(config)#bfd
huawei(config-bfd)#quit
huawei(config)#bfd ToRouter_1 bind peer-ip 10.10.10.2 interface vlanif 30
huawei(config-bfd-session-torouter_1)#discriminator local 1
huawei(config-bfd-session-torouter_1)#discriminator remote 2
huawei(config-bfd-session-torouter_1)#commit
huawei(config-bfd-session-torouter_1)#quit
huawei(config)#interface vlanif 30
huawei(config-if-vlanif30)#isis bfd static
huawei(config-if-vlanif30)#quit
huawei(config)#bfd ToRouter_2 bind peer-ip 20.20.20.2 interface vlanif 40
huawei(config-bfd-session-torouter_2)#discriminator local 3
huawei(config-bfd-session-torouter_2)#discriminator remote 4
huawei(config-bfd-session-torouter_2)#commit
huawei(config-bfd-session-torouter_2)#quit
huawei(config)#interface vlanif 40
huawei(config-if-vlanif40)#isis bfd static
huawei(config-if-vlanif40)#quit
```

Step 4 Configure dynamic BFD for IS-IS.

```
huawei(config)#isis
huawei(config-isis-1)bfd all-interfaces enable
huawei(config-isis-1)#quit
huawei(config)#interface vlanif 30
huawei(config-if-vlanif30)#isis bfd enable
huawei(config-if-vlanif30)#isis bfd min-tx-interval 10 min-rx-interval 10
detect-multiplier 3
huawei(config-if-vlanif30)#quit
huawei(config)#interface vlanif 40
huawei(config-if-vlanif40)#isis bfd enable
huawei(config-if-vlanif40)#isis bfd min-tx-interval 10 min-rx-interval 10
detect-multiplier 3
huawei(config-if-vlanif40)#quit
```

Step 5 Save the data.

```
huawei(config)#save
```

----End

Result

After establishing the neighbor relation with each router through IS-IS, the access node creates two BFD sessions. When the active link is faulty, its bound BFD session is down, which triggers the IS-IS neighbor relation to be down. Therefore, the route is switched to the standby link.

Run the **display isis bfd interface** command to query the BFD configuration information of an IS-IS interface.

Run the **display isis bfd session** command to query the BFD session information.

Configuring BFD for BGP or BGP4+

BFD for BGP or BGP4+ speeds up fault detection and therefore increases the route convergence speed.

Prerequisites

Basic BGP or BGP4+ functions are configured.

Context

BGP periodically sends Keepalive packets to its peers to detect the status of its peers. The detection mechanism, however, takes more than one second. When the data transmission rate reaches the level of Gbit/s, such slow detection will cause a large amount of data to be lost. As a result, the requirement for high reliability of carrier-class networks cannot be met.

BFD for BGP or BGP4+ detects faults on links between BGP peers within 50 milliseconds. The fast detection speed ensures fast BGP route convergence and minimizes traffic loss.

NOTE

By default, a multi-hop IBGP session is established between Huawei devices that set up an IBGP peer relationship. A BFD for IGP session and a BFD for IBGP session cannot be both set up between a Huawei device and a non-Huawei device that sets up a single-hop BGP session with its peer by default. In such a situation, setting up only a BFD for IGP session or a BFD for IBGP session between the Huawei and non-Huawei devices is recommended.

Procedure

Enable BFD globally.

1. Run the **bfd** command to enable BFD globally and enter the BFD mode.

   BFD must be enabled globally before configurations relevant to BFD are performed. By default, BFD is disabled globally.
2. Run the **quit** command to quit the BFD mode.

Step 1 Create a BFD session.

1. Run the **bgp** command to enter the BGP mode.
2. (Optional) Configure BFD for BGP or BGP4+ in the VPN. To configure BFD for BGP or BGP4+ for the public network, skip this step.
   - To configure BFD for BGP, run the **ipv4-family vpn-instance** *vpn-instance-name* command to enter the BGP-VPN instance IPv4 address family mode.

- To configure BFD for BGP4+, run the **ipv6-family vpn-instance** *vpn-instance-name* command to enter the BGP-VPN instance IPv6 address family mode.

3. Run the **peer bfd enable** command to enable BFD for the peer or peer group, and a BFD session is established.

   A BFD session is set up only when the BGP session is in the Established state.

   After BFD is enabled for a peer group, BFD sessions will be created on the peers that belong to this peer group and are not configured with the **peer bfd block** command.

Step 2 (Optional) Run the **peer bfd** command to configure the BFD parameters.

The BFD parameters of peers take precedence over those of peer groups. If BFD parameters are configured on peers, they will be used in BFD session establishment.

The default interval for transmitting BFD packets and the default detection multiplier are recommended. When changing the default values, pay attention to the network status and the network reliability requirement. A short interval for transmitting BFD packets can be configured for a link that has a higher reliability requirement. A long interval for transmitting BFD packets can be configured for a link that has a lower reliability requirement. There are three formulas:

- Actual interval for the local device to send BFD packets = max {Locally configured interval for transmitting BFD packets, Remotely configured interval for receiving BFD packets}
- Actual interval for the local device to receive BFD packets = max {Remotely configured interval for transmitting BFD packets, Locally configured interval for receiving BFD packets}
- Local detection period = Actual interval for receiving BFD packets x Remotely configured BFD detection multiplier

For example:

- On the local device, the configured interval for transmitting BFD packets is 200 ms, the interval for receiving BFD packets is 300 ms, and the detection multiplier is 4.
- On the peer device, the configured interval for transmitting BFD packets is 100 ms, the interval for receiving BFD packets is 600 ms, and the detection multiplier is 5.

Then:

- On the local device, the actual interval for transmitting BFD packets is 600 ms calculated by using the formula max {200 ms, 600 ms}; the interval for receiving BFD packets is 300 ms calculated by using the formula max {100 ms, 300 ms}; the detection period is 1500 ms calculated by multiplying 300 ms by 5.
- On the peer device, the actual interval for transmitting BFD packets is 300 ms calculated by using the formula max {100 ms, 300 ms}; the interval for receiving BFD packets is 600 ms calculated by using the formula max {200 ms, 600 ms}; the detection period is 2400 ms calculated by multiplying 600 ms by 4.

wtr *wtr-value* can be specified in the command to suppress frequent BFD and BGP session flapping caused by link flapping. If a BFD session over a link goes Down, it does not go Up immediately after the link recovers. Instead, the BFD session waits for the WTR timer to expire before going Up. If the link fails again before the WTR timer expires, BFD does not send a link fault message to BGP, and the BGP session status is stabilized.

Step 3 (Optional) Run the **peer bfd block** command to prevent a peer from inheriting the BFD function of the peer group to which it belongs.

If a peer joins a peer group enabled with BFD, the peer inherits the BFD configuration of the group and creates a BFD session. To prevent the peer from inheriting the BFD function of the peer group, perform this step.

The **peer bfd block** command and the **peer bfd enable** command are mutually exclusive. After the **peer bfd block** command is run, the BFD session is automatically deleted.

Step 4 Query the BFD session information.

- Run the **display bgp bfd session** command to query the BFD session between BGP peers.
- Run the **display bgp ipv6 bfd session** command to query the BFD session between BGP4+ peers.

----End

Configuring BFD for RSVP

BFD for RSVP is applied to a scenario where TE FRR is used and a Layer 2 device exists on the primary LSP between a PLR and its downstream neighbors. On a network where GR is enabled on the PLR and MP, BFD for RSVP is also recommended.

Prerequisites

The RSVP-TE tunnel is configured.

Context

By default, the interval at which RSVP Hello messages are sent is 3 seconds. The interval at which a neighbor is declared Down is three times the interval at which RSVP Hello messages are sent. This allows devices to detect a fault in an RSVP neighbor at seconds level. If a Layer 2 device exists on a link between RSVP neighboring nodes, the neighboring node cannot rapidly detect the fault after the link fails, resulting in a great loss of data.

BFD detects faults at millisecond level in protected links or nodes. BFD for RSVP rapidly detects faults in an RSVP neighbor, allowing packets to switch to a backup LSP rapidly.

NOTE

BFD for LSP can function properly though the forward path is an LSP and the backward path is an IP link. The forward path and the backward path must be established over the same link; otherwise, if a fault occurs, BFD cannot identify the faulty path. Before deploying BFD, ensure that the forward and backward paths are over the same link so that BFD can correctly identify the faulty path.

Perform the following steps on the two RSVP neighboring nodes between which a Layer 2 device resides:

Procedure

Enable BFD globally.

1. Run the **bfd** command to enable BFD globally and enter the BFD mode.

   BFD must be enabled globally before configurations relevant to BFD are performed. By default, BFD is disabled globally.
2. Run the **quit** command to quit the BFD mode.

Step 1 Create a BFD session.

- If most RSVP interfaces on a node need BFD for RSVP, enable BFD for RSVP globally.

1. Run the **mpls** command to enter the MPLS mode.
2. Run the **mpls rsvp-te bfd all-interfaces enable** command to enable BFD for RSVP globally.

   After this command is run in the MPLS mode, BFD for RSVP is enabled on all RSVP interfaces except the interfaces with BFD for RSVP that are blocked.
3. (Optional) Block BFD for RSVP on the RSVP interfaces that need not BFD for RSVP.

   Run the **interface vlanif** command to enter the VLANIF interfac mode, and run the **mpls rsvp-te bfd block** command to block BFD for RSVP.

- If certain RSVP interfaces on a node need BFD for RSVP, enable BFD for RSVP on the RSVP interfaces.

1. Run the **interface vlanif** command to enter the VLANIF interface mode.
2. Run the **mpls rsvp-te bfd enable** command to enable BFD for RSVP on an RSVP interface.

Step 2 (Optional) Configure the BFD parameters.

- If most RSVP interfaces on a node use the same BFD parameters, configure global BFD parameters.

1. Run the **mpls** command to enter the MPLS mode.
2. Run the **mpls rsvp-te bfd all-interfaces** command to configure global BFD parameters.

- If certain RSVP interfaces require BFD parameters different from global BFD parameters, configure BFD parameters on the RSVP interfaces.

1. Run the **interface vlanif** command to enter the VLANIF interface mode.
2. Run the **mpls rsvp-te bfd** command to configure BFD parameters on an RSVP interface.

Step 3 Run the **display mpls rsvp-te bfd session** command to query the BFD session information.

----End

19.10.5 References

The following table lists the references.

Document No.	Document Name	Protocol Compliance
RFC 5880	Bidirectional Forwarding Detection	• Compliant only with the synchronous mode • Compliant except for authentication
RFC 5881	BFD for IPv4 and IPv6 (Single Hop)	Fully compliant
RFC 5882	Generic Application of BFD	• Compliant except for OSPF virtual connection
RFC 5883	BFD for Multihop Paths	• Compliant except for discriminator learning in outband mode • Compliant except for

Document No.	Document Name	Protocol Compliance
		bidirectional connection
RFC 5885	BFD for PW	Compliant only with IP/UDP encapsulation

19.11 Ring Check

The ring check feature is mainly used to detect and eliminate the user-side ring network.

19.11.1 Introduction

Definition

Ring check is a function of detecting the ring network formed on the user side. The ring check feature enables the device to transmit the ring check packets to the user port periodically, and to monitor the ring check packets received on the user side and the network side to check whether a loop occurs on the network of the carrier. If a loop occurs, the MA5600T/MA5603T/MA5608T deactivates the user ports on the loop and reports the corresponding alarm to the NMS. This ensures that the device runs in the normal state and that the services of other users are not affected.

Purpose

Ring check is used to quickly locate the user-side ring network, and eliminate the ring network according to requirements.

- To prevent the self-loop on a single user port from occurring
- To prevent the loop between user ports from occurring
- To prevent the loop between a user port and a network port from occurring

Benefit

Benefits to carriers

The Ring check feature enables the system to detect the carrier's network and report an alarm to the NMS when a loop occurs. The alarm enables the carrier to know the network fault in the shortest period of time so that the fault can be quickly rectified to resume the normal running of the network.

Benefits to users

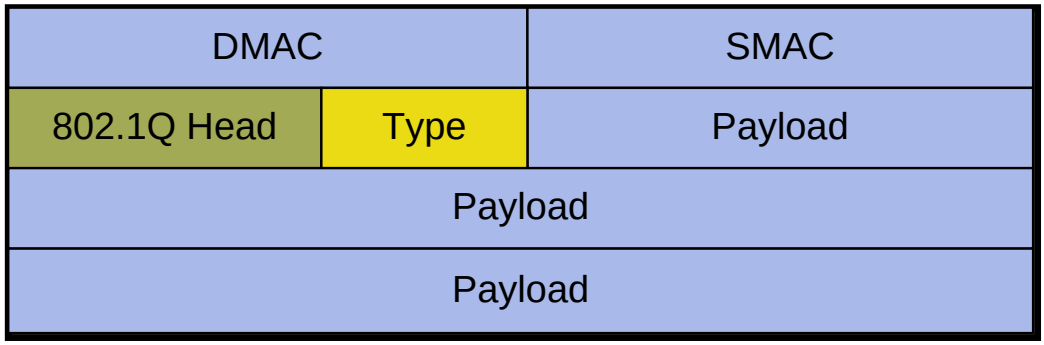
The Ring check feature enables the device to deactivate the user port on a loop to ensure that authorized users receive a good network service rather than be affected.

19.11.2 Principle

Format of the Ring Check Packet

The ring check feature enables the device to transmit the ring check packets to the user port periodically, and to monitor the ring check packets received on the user side and the network side to check whether a loop occurs on the network of the carrier. Figure 19-73 shows the format of the ring check packet.

Figure 19-73 Format of the ring check packet



- DMAC indicates the broadcast MAC address with value 0xFF and SMAC indicates the bridge MAC address.
- 802.1Q Head is optional according to flow attributes on the user side.
- Type indicates the proprietary Ethernet type, which can be configured.
- Payload of the packet content is proprietary and it needs not be configured.

Principle

After the ring check function is enabled, the device periodically transmits private ring check packets to the user port and captures the user-side ring check packets on the network and user sides simultaneously.

- As for the ring check packets captured on the network side, the system first checks whether they are transmitted from the local device.
  - If yes, the system finds out the source port transiting the ring check packets and reports an alarm to the NMS, but does not deactivate this source port. This is because a user can forge the ring check packets and the system cannot determine whether the ring check packets are forged by a user or are transmitted from the device. The check performed by the system prevents misjudgment of the check point.
  - If not, the system discards the packets.
- As for the ring check packets captured on the user side, the system reports an alarm to the NMS and deactivates the port receiving the packets, thus eliminating the loop in the network.

NOTE

After the fault is solved, the port will be restarted after some time. If you want to enable the port quickly, the port needs to be deactivated and activated again.

- Detects a maximum of 12 up traffic streams per second. If the system has 8K up traffic steams, 682.67s (8192/12) is required for detecting a loop if a loop exists.

Figure 19-74 shows the use-side ring network scenarios in FTTH/DSLAM applications.

- In the case of (1), (2), (3), and (4), as for the ring check packets captured on the user side, the system directly deactivates the port receiving the packets, thus eliminating the loop in the network.
- In the case of (4), this kind of network topology needs to be prevented. This is because the system cannot determine whether the ring check packets captured on the network side are forged by a user or are transmitted from the device. This kind of network topology prevents misjudgment of the check point.

Figure 19-74 Use-side ring network scenarios in FTTH/DSLAM applications

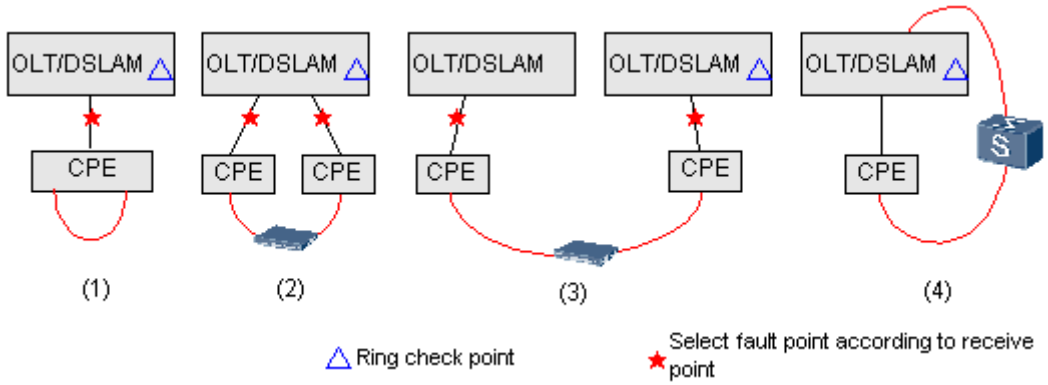
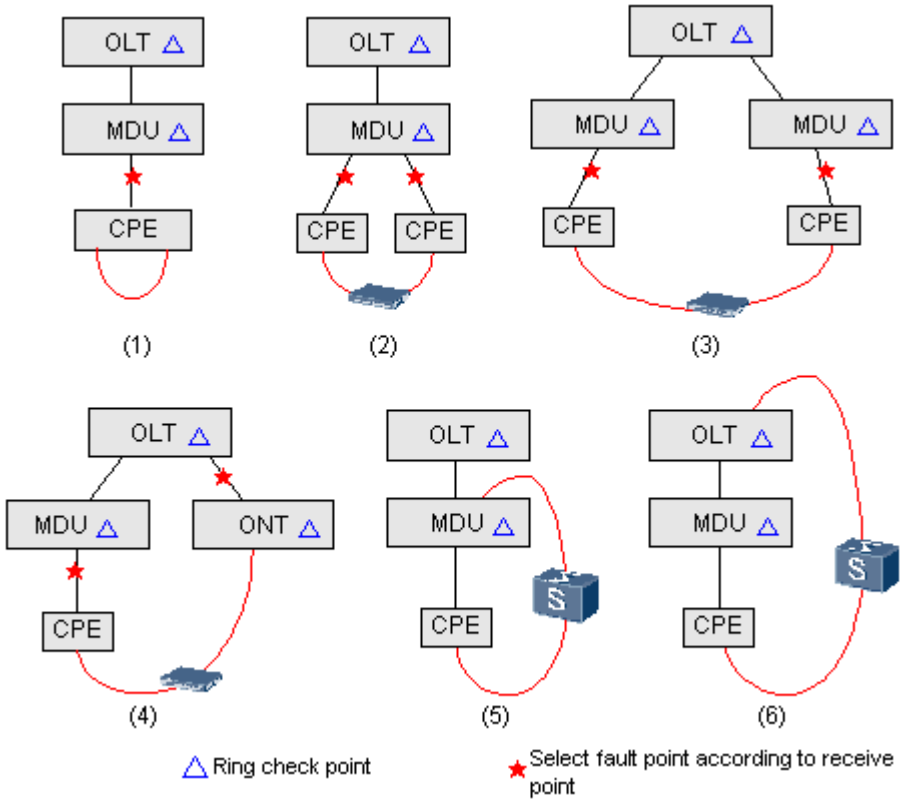


Figure 19-75 shows the use-side ring network scenarios in FTTB/FTTC applications.

- The Ethernet Types of ring check packets of the OLT and the MDU are recommended to be the same.
  - If Ethernet Types of ring check packets of the OLT and the MDU are the same, the ring check packets are captured on the OLT and the MDU and are judged. In the case of (6), the ring check packets transmitted by the MDU are terminated by the OLT. Therefore, ring network cannot be detected in this kind of network topology.
  - If Ethernet Types of ring check packets of the OLT and the MDU are different, the OLT and the MDU capture its own ring check packets. In the case of (4), the MDU and the ONT connected to the OLT are not interconnected. Therefore, ring network cannot be detected in this kind of network topology.
- In the case of (1), (2), (3), and (4), as for the ring check packets captured on the user side, the system directly deactivates the port receiving the packets, thus eliminating the loop in the network. In the case of (5), the system can detect the ring network but not deactivate the port to eliminate the ring network, instead, the system reports an alarm.
- In the case of (5) and (6), these two kinds of network topology need to be prevented. This is because the system cannot determine whether the ring check packets captured on the network side are forged by a user or are transmitted from the device. This kind of network topology prevents misjudgment of the check point.

Figure 19-75 Use-side ring network scenarios in FTTB/FTTC applications



19.11.3 Configuring the Ring Network Detection on the User Side

This topic describes how to configure the ring network detection on the user side and network side to prevent the services from being affected by the ring network.

Context

- By default, the ring network detection is disabled.
- After the ring network detection is enabled, the system automatically detects the ring network.

NOTICE

To ensure the security of the device, it is recommended that the ring network detection is enabled.

Procedure

Run the **ring check enable** command to enable the ring network detection.

Step 1 (Option) Run the **ring check private-ethtype** command to configure the private Ethernet protocol type of ring network detection packets. If the default Ethernet protocol type of ring network detection packets is the same as the existing protocol type on the network, change the

protocol type of the ring detection packets to be different from the protocol type on the network.

Step 2 (Option) Run the **ring check resume-interval** command to set the auto-activation interval for the port in ring check.

After the auto-activation interval is set, when detecting a ring network, the system automatically activates the port that sends the ring check packet.

Step 3 Run the **display ring check config** command to query the status of the ring network detection.

----End

Example

To enable the ring network detection, do as follows:

```
huawei(config)#ring check enable
huawei(config)#display ring check config
  Ring check switch status   : enable
  Ring check private ethtype : 0x8300
  Ring check resume-interval(min) : -
```

20 NE Cascading

About This Chapter

NE cascading refers to a networking mode in which the MA5600T/MA5603T/MA5608T series are directly connected to each other through the FE/GE port on the board. NE cascading saves the upstream optical fiber resources of the access node.

The MA5600T/MA5603T/MA5608Ts (NEs) can be directly connected to each other through the FE or GE port. Cascading saves the uplink optical fibers and simplifies networking and service configuration.

20.1 Introduction to NE Cascading

Definition

NE cascading refers to a networking mode in which the MA5600T/MA5603T/MA5608T series are directly connected to each other through the FE/GE port on the board.

Purpose

NE cascading makes the networking of the MA5600T/MA5603T/MA5608Ts more flexible, and saves the upstream optical fiber resources of the access node. In addition, remote subtending saves the convergence devices at the central office (CO), simplifies topology, and facilitates service configuration.

20.2 NE Cascading Principle

According to the device location, subtending supported by the MA5600T/MA5603T/MA5608T can be local subtending or remote subtending.

Local Subtending

Local subtending refers to the subtending of multiple MA5600T/MA5603T/MA5608T subracks that are in a cabinet or in multiple local cabinets.

The local subtending of MA5600T/MA5603T/MA5608Ts can be implemented through the control board or GIU upstream board. Each GIU upstream board provides up to four GE optical ports for upstream transmission or subtending. The number of the ports for subtending depends on the bandwidth requirement. If an active/standby configuration is required, configure two GIU upstream boards.

- According to the connection type, local subtending can be in a star topology or in a daisy chain topology.
- According to the configuration of the two GIU upstream boards, local subtending can be implemented through a single GIU upstream board or through dual GIU upstream boards.
- The local subtending in a star topology is shown in Figure 20-1 and Figure 20-2. The local subtending in a daisy chain topology is shown in Figure 20-3 and Figure 20-4.

Figure 20-1 Local subtending in a star topology (MA5600T)

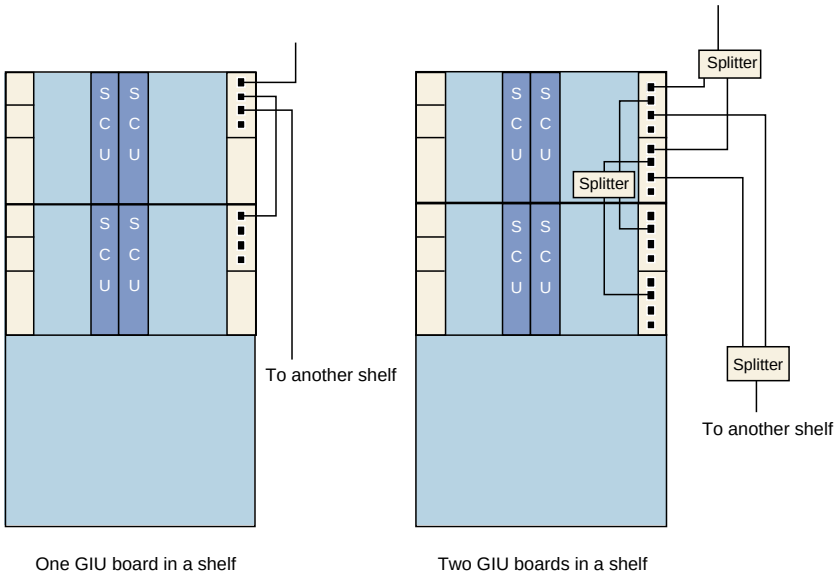
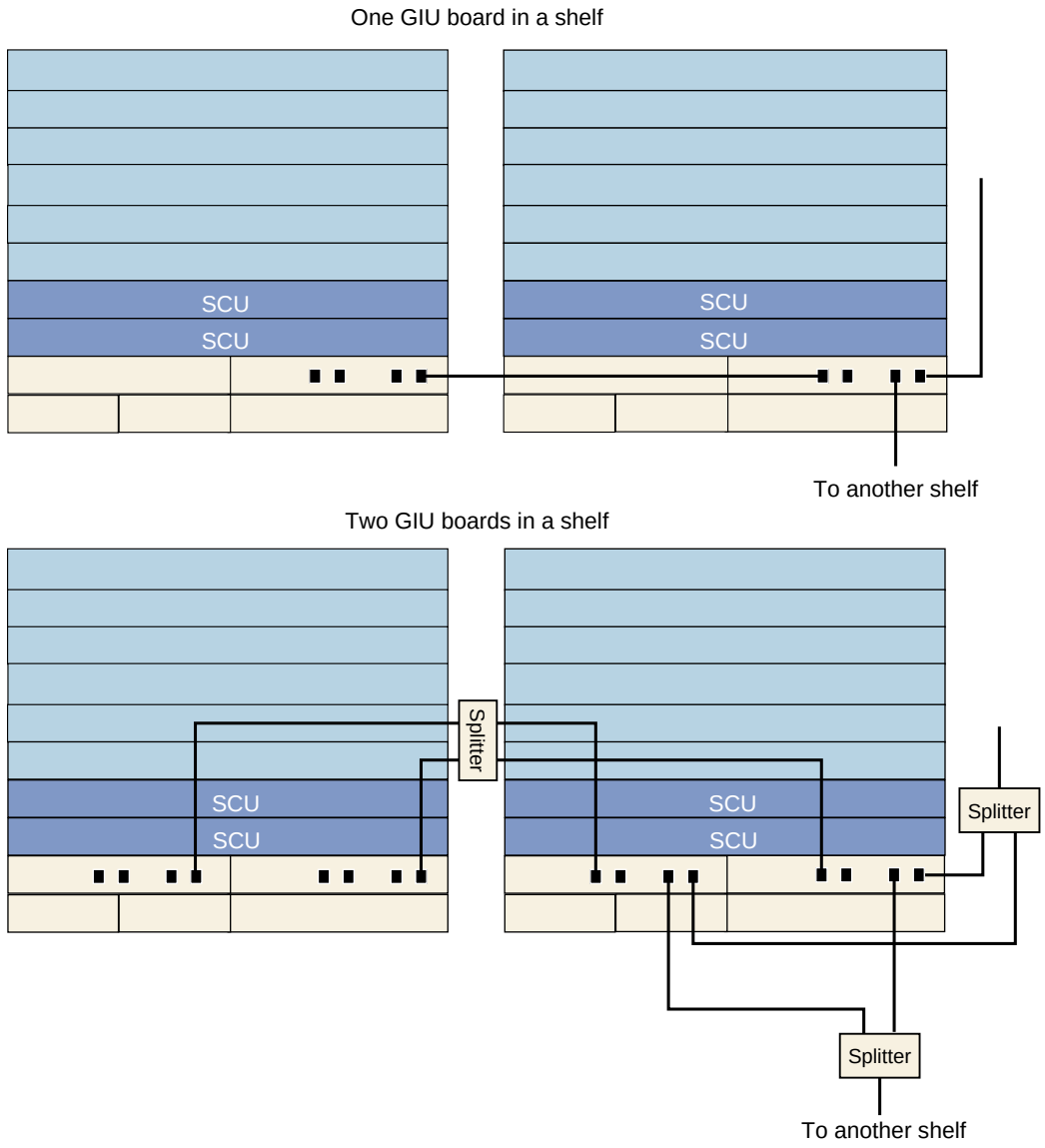


Figure 20-2 Local subtending in a star topology (MA5603T)



NOTE

Optical splitters are required for local subtending in the case of the configuration of dual GIU upstream boards.

Figure 20-3 Local subtending in a daisy chain topology (MA5600T)

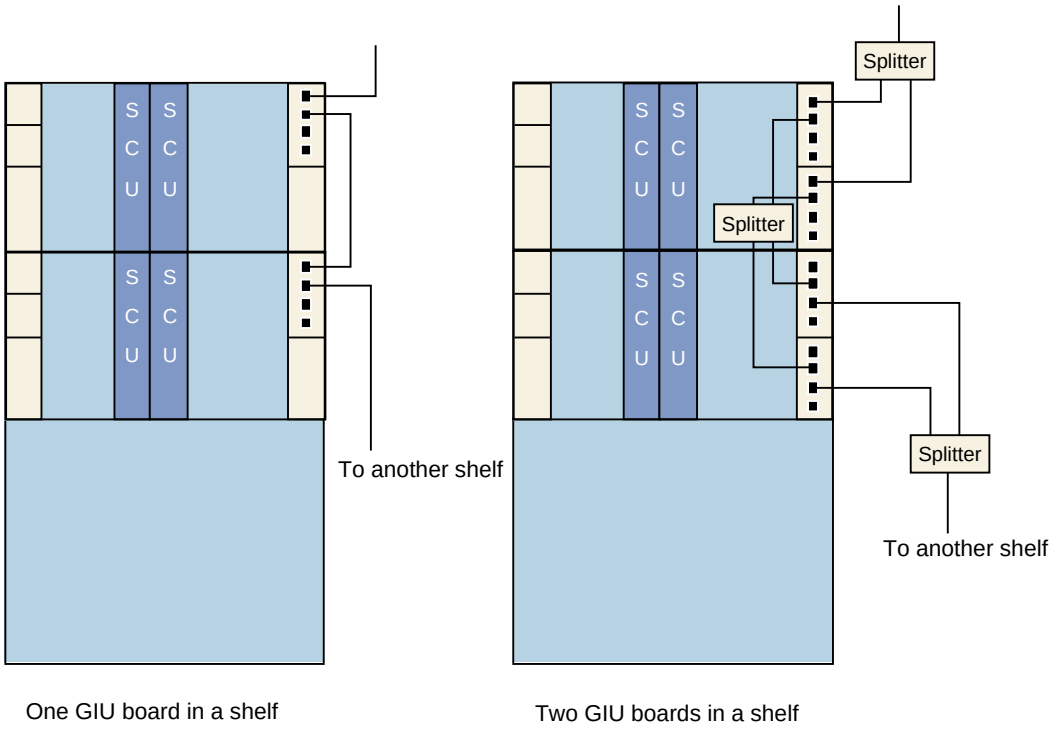
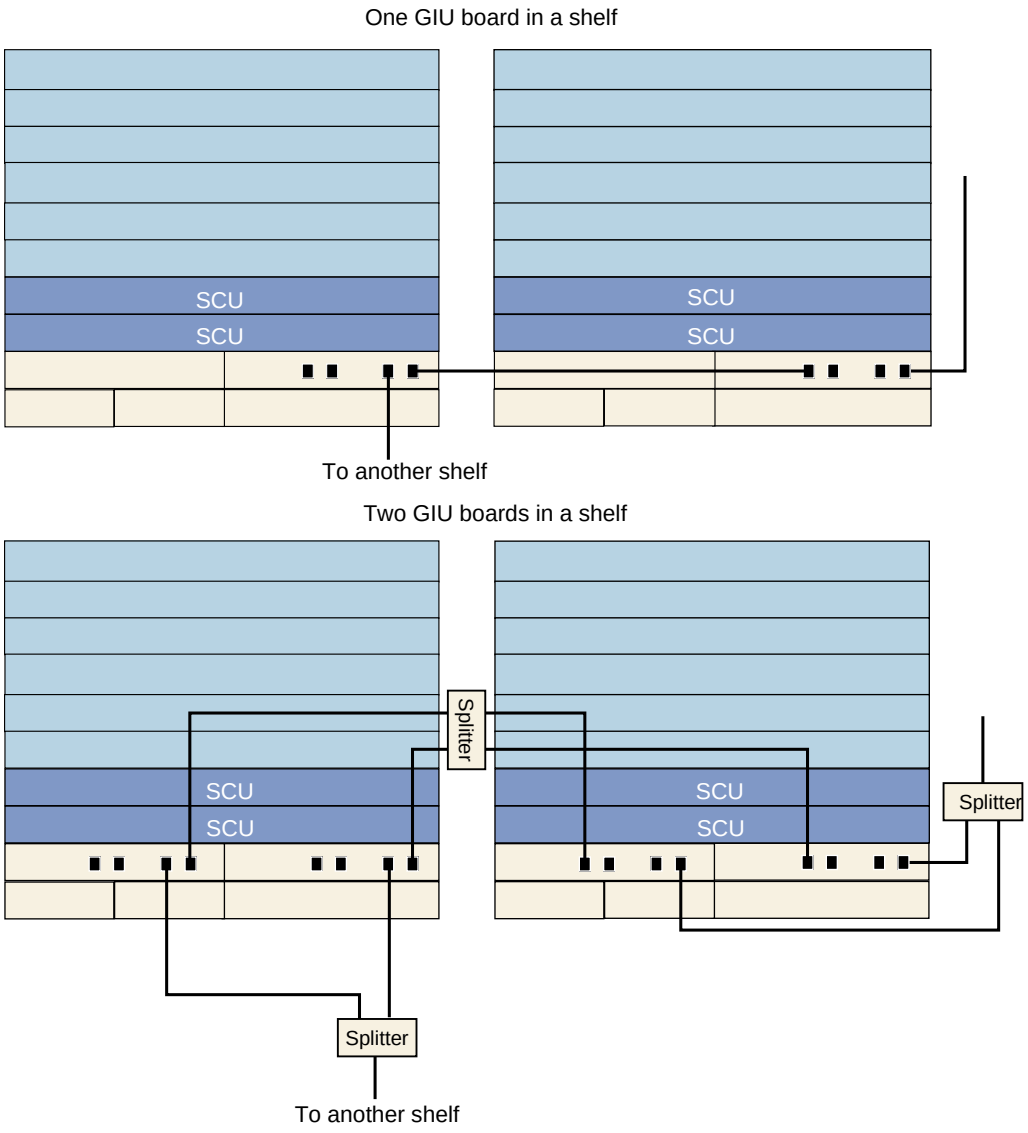


Figure 20-4 Local subtending in a daisy chain topology (MA5603T)



Remote Subtending with a Control Device

Remote subtending (with a control device) refers to the subtending mode in which the control MA5600T/MA5603T/MA5608T subtends the remote MA5600T/MA5603T/MA5608T or other mini DSLAMs such as the MA5606T through optical fibers.

In remote subtending (with the control device), the port on the control board or GIU board can be used for subtending. The control board and GIU board support up to eight ports. If the number of subtending ports is insufficient, an additional ETHA or ETHB board can be added to provide subtending ports. For example, as shown in Figure 20-5and Figure 20-6, the ETHA or ETHB board is used for subtending the remote MA5600T/MA5603T/MA5608T.

Figure 20-5 Remote subtending with a control device (MA5600T)

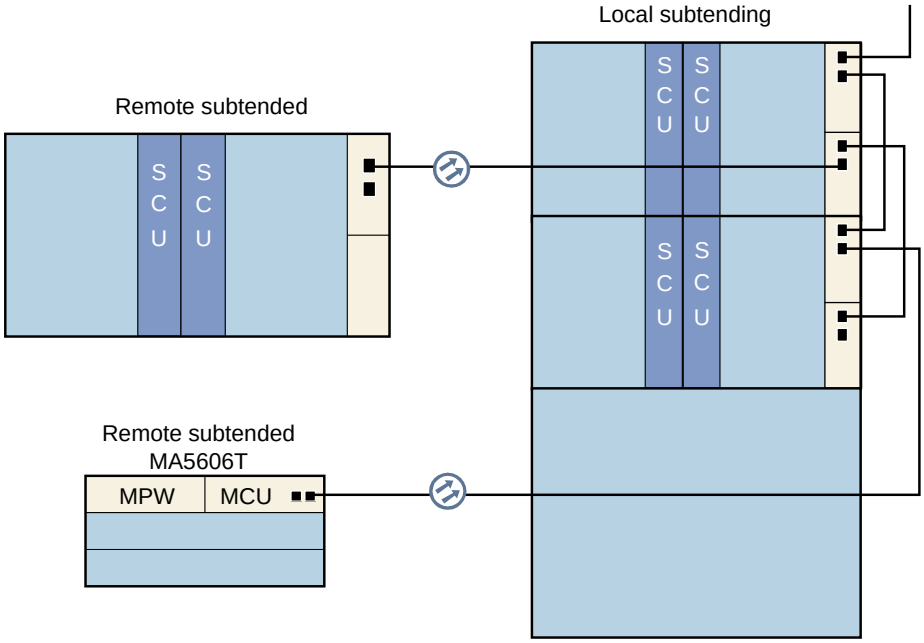
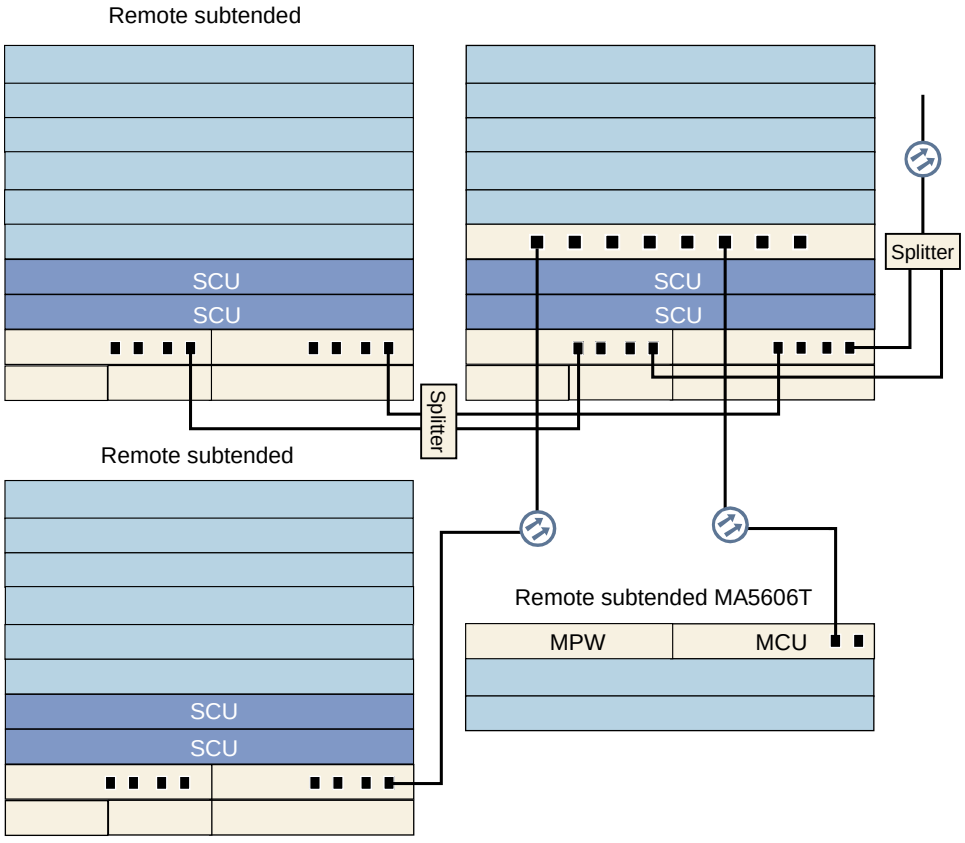


Figure 20-6 Remote subtending with a control device (MA5603T)



When the devices have the control module, remote subtending can also be implemented in an MSTP/RSTP ring topology, as shown in Figure 20-7 and Figure 20-8. In this application, the control board, the GIU board, and the ETHA or ETHB subtending board all support the subtending in an MSTP ring topology; however, the ETHA or ETHB board can be configured only on the device that is connected to the CO device and provides upstream ports. A more complicated subtending network is an MSTP/RSTP ring network in which each node is involved in local or remote subtending.

Figure 20-7 Remote subtending in an MSTP/RSTP ring topology (MA5600T)

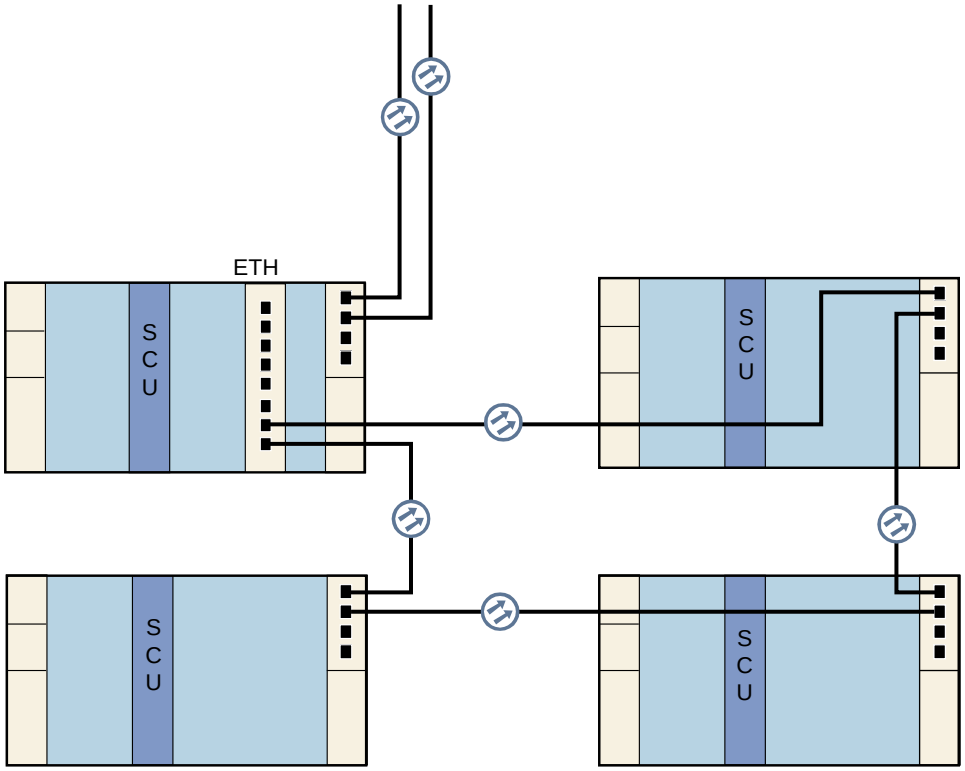
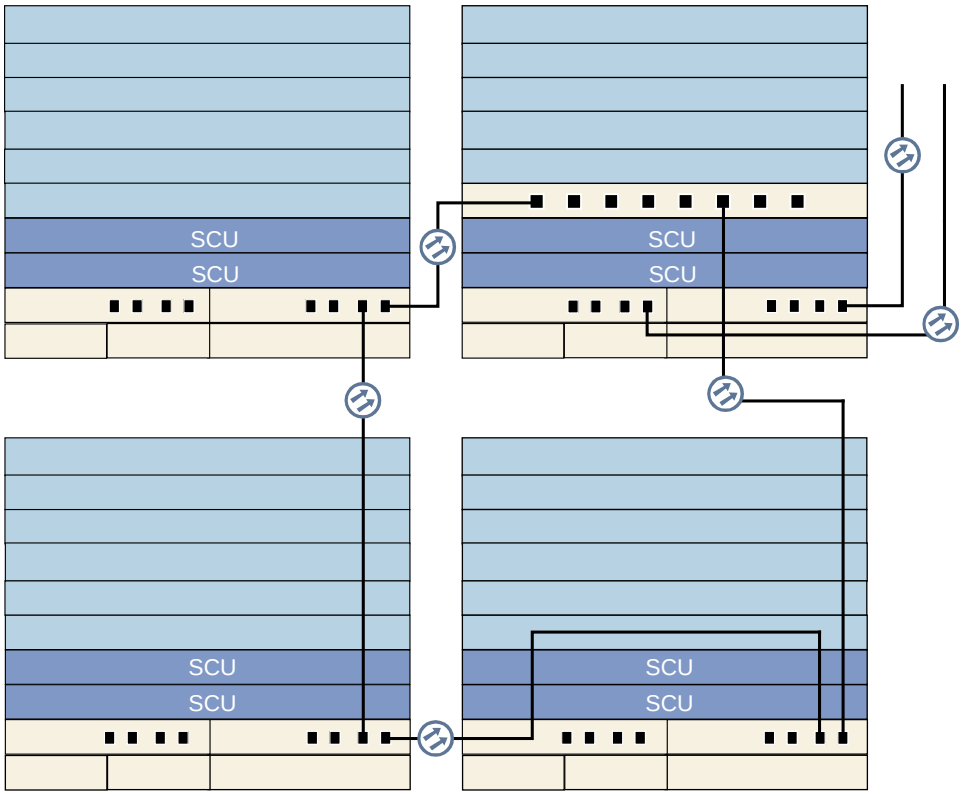


Figure 20-8 Remote subtending in an MSTP/RSTP ring topology (MA5603T)



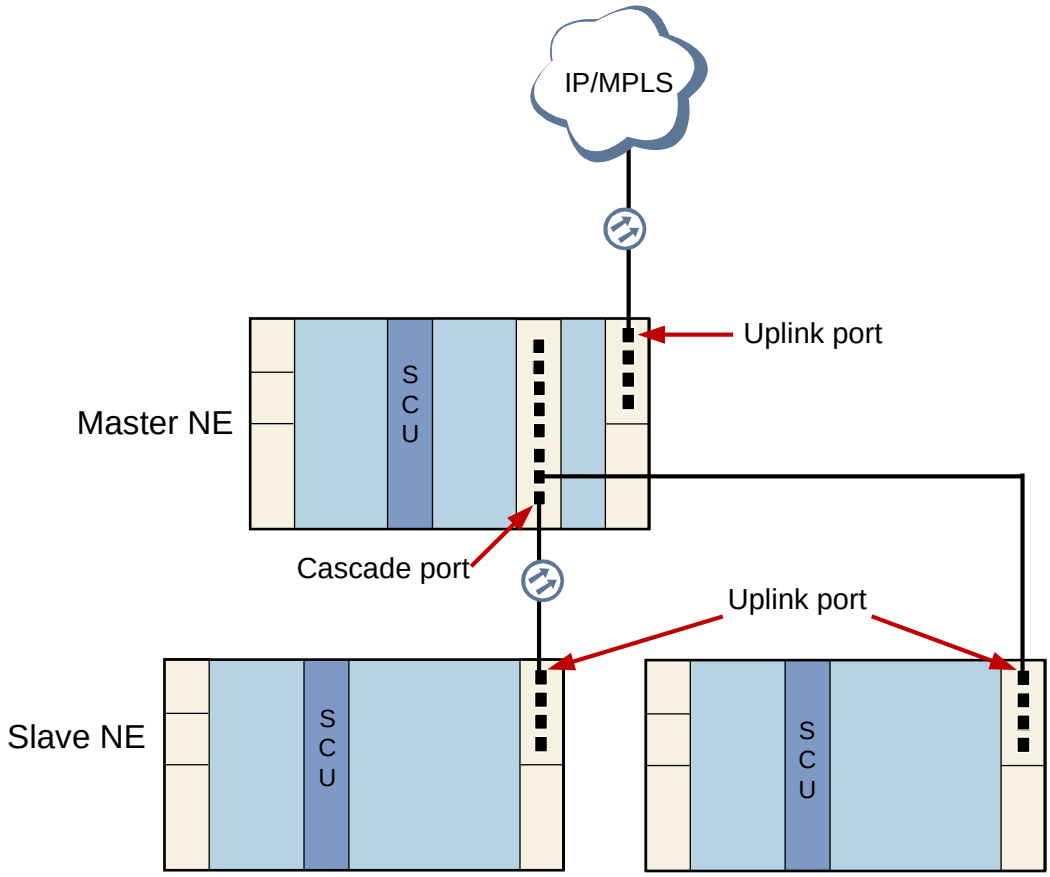
20.3 Configuring NE Cascade and Uplink Transmission Through the FE or GE Port

The MA5600T/MA5603T/MA5608Ts (NEs) can be directly connected to each other through the FE or GE port. Cascading saves the uplink optical fibers and simplifies networking and service configuration.

Context

The cascade ports on master NEs and uplink ports on slave NEs must have the same port type, port rate, and port duplex mode.

Figure 20-9 Cascading network



NOTE

Configure the network role of a board first. Otherwise, the network role may fail to be switched if other configurations have been performed.

Table 20-1 Configuration comparison of cascade and uplink boards

Board	Configure the Network Role of a Board		Add the Uplink or Cascade Port into a VLAN
	Enter the board mode	Configure the network role of a board	
GIU board	Run the **interface giu** command to enter the GIU mode.	Based on port • uplink • cascade Default: uplink	Run the **port vlan** command.
SCU control board	Run the **interface scu** command to enter the SCU mode.		Run the **port vlan** command.
MCU control board	Run the **interface mcu** command to enter the MCU mode.		Run the **port vlan** command.

Board	Configure the Network Role of a Board		Add the Uplink or Cascade Port into a VLAN
	Enter the board mode	Configure the network role of a board	
ETHB	Run the **interface eth** command to enter the ETH mode.	Based on board • uplink • cascade • extend Default: cascade	Run the **port vlan** command.
SPUA/SPUC	Run the **interface eth** command to enter the ETH mode.	Based on board • uplink • cascade Default: uplink	Run the **service-port** command to create an Ethernet service flow to add the uplink or cascade port into a VLAN. If multiple uplink ports are added into a protection group, these ports must be added into an aggregation group. In this way, you only need to create a service flow to add the master port of the protection group into the VLAN so that all the member ports of the protection group will be added into the VLAN.
SPUF	Run the **interface eth** command to enter the ETH mode.		Run the **port vlan** command.
OPGD/OPGE	Run the **interface opg** command to enter the OPG mode.	Based on board (uplink is not supported) • cascade • user Default: user	Run the **service-port** command.

Procedure

- Configure the master NE
  a. Set the network role of the uplink board or port of the master NE.
     Enter the specific board mode according to the uplink board types. Run the **network-role** [*port-id*] **uplink** command to set the network role of the board or port to uplink.
  b. Set the network role of the cascade board or port of the master NE.

Enter the specific board mode according to the cascade board types. Run the **network-role** [*port-id*] **cascade** command to set the network role of the board or port to cascading.

c. Configure the VLAN of the master NE.

The VLAN type is smart, and the VLAN attribute is common. For details about the configuration, see 13.3.9 Configuring a VLAN.

d. Add an uplink port to the VLAN of the master NE.

According to the uplink board types, run the **port vlan** or **service-port** command in global config mode to add the uplink port to the VLAN.

e. Add a cascade port to the VLAN of the master NE.

According to the cascade board types, run the **port vlan** or **service-port** command in global config mode to add the cascade port to the VLAN.

- Configure the slave NE

a. Set the network role of the uplink board or port of the slave NE.

Enter the specific board mode according to the uplink board types. Run the **network-role** [*port-id*] **uplink** command to set the network role of the board or port to uplink.

b. Configure the VLAN of the slave NE. The VLAN of the slave NE is the same as the VLAN of the master VLAN.

The VLAN type is smart, and the VLAN attribute is common. For details about the configuration, see 13.3.9 Configuring a VLAN.

c. Add an uplink port to the VLAN of the slave NE.

According to the uplink board types, run the **port vlan** or **service-port** command in global config mode to add the uplink port to the VLAN.

----End

Example

Assume that master NE huawei_A and slave NE huawei_B are cascaded through the GIU board. To add uplink port 0/19/0 and cascade port 0/19/1 of huawei_A to VLAN 100, and add uplink port 0/19/0 of huawei_B to VLAN 100, do as follows:

```
huawei_A(config)#interface giu 0/19
huawei_A(config-if-giu-0/19)#network-role 1 cascade
huawei_A(config-if-giu-0/19)#quit
huawei_A(config)#vlan 100 smart
huawei_A(config)#port vlan 100 0/19 0
huawei_A(config)#port vlan 100 0/19 1

huawei_B(config)#vlan 100 smart
huawei_B(config)#port vlan 100 0/19 0
```

Assume that master NE huawei_A and slave NE huawei_B are cascaded through the ETHB board. To add uplink port 0/19/0 and cascade port 0/2/0 of huawei_A to VLAN 100, and add uplink port 0/19/0 of huawei_B to VLAN 100, do as follows:

```
huawei_A(config)#interface eth 0/2
huawei_A(config-if-eth-0/2)#network-role cascade
huawei_A(config-if-eth-0/2)#quit
huawei_A(config)#vlan 100 smart
```

```
huawei_A(config)#port vlan 100 0/19 0
huawei_A(config)#port vlan 100 0/2 0

huawei_B(config)#vlan 100 smart
huawei_B(config)#port vlan 100 0/19 0
```

Assume that master NE huawei_A and slave NE huawei_B are cascaded through the SPUA board. To add uplink port 0/19/0 and cascade port 0/2/0 of huawei_A to VLAN 100, and add uplink port 0/2/1 on the SPUA board of huawei_B to VLAN 100, do as follows:

```
huawei_A(config)#interface eth 0/2
huawei_A(config-if-eth-0/2)#network-role cascade
huawei_A(config-if-eth-0/2)#quit
huawei_A(config)#vlan 100 smart
huawei_A(config)#port vlan 100 0/19 0
//Configure cascading on the SPUA board.
huawei_A(config)#service-port vlan 100 eth 0/2/0 multi-service user-vlan 100

huawei_B(config)#vlan 100 smart
//Configure upstream transmission on the SPUA board.
huawei_B(config)#service-port vlan 100 eth 0/2/1 multi-service user-vlan 100
```

20.4 NE Cascading Reference Standards and Protocols

The following is the reference standard of this feature:

IEEE 802.1w Rapid Spanning Tree Protocol

21 Remote Software Commissioning (GE)

About This Chapter

This section describes the implementation principles and configuration of remote software commissioning using GE upstream transmission.

21.1 Introduction

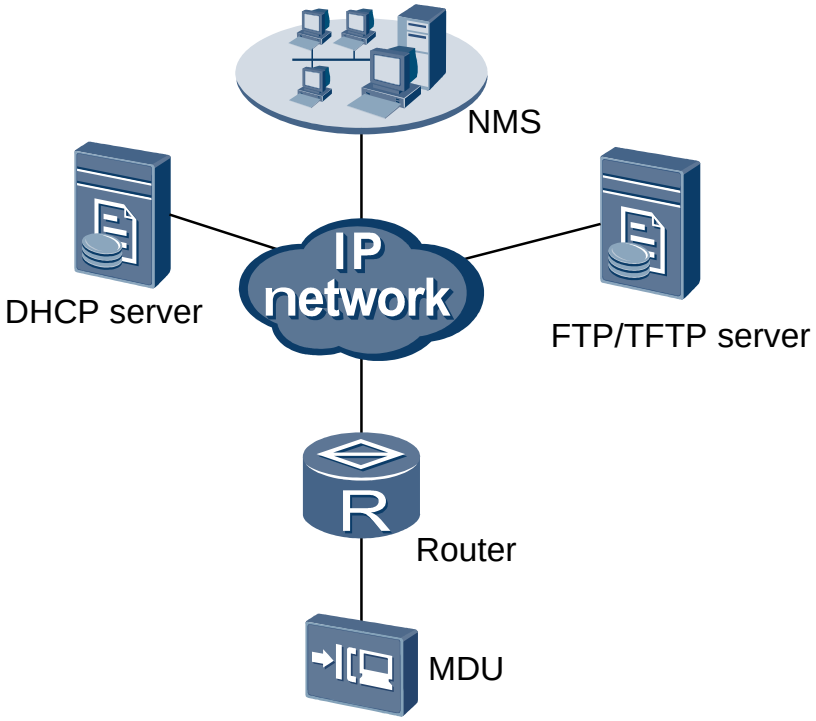
During site deployment for a multi-dwelling unit (MDU) using GE upstream transmission, or an MA5800 using a GE port to cascade an MA5600T/MA5603T/MA5608T, the NMS can manage the MDU/MA5600T/MA5603T/MA5608T only after installation engineers install the MDU/MA5600T/MA5603T/MA5608T and commissioning engineers configure its network management channel parameters onsite. The MDU/MA5600T/MA5603T/MA5608T supports remote software commissioning. This allows the NMS to manage the MDU/MA5600T/MA5603T/MA5608T without commissioning engineers having to first configure network management channel parameters onsite, and thereby reduces site deployment costs. After the MA5600T/MA5603T/MA5608T is powered on, it automatically configures management parameters. After the MDU is powered on, it automatically configures management and service data.

Remote software commissioning using GE upstream transmission can be implemented based on Dynamic Host Configuration Protocol (DHCP) or neighbor automatic communication (NAC).

DHCP-based Remote Software Commissioning Using GE Upstream Transmission

Figure 21-1 shows the networking of DHCP-based remote software commissioning using GE upstream transmission.

Figure 21-1 Networking of DHCP-based remote software commissioning using GE upstream transmission



1. The IP and MAC addresses of the cascaded device are bound on the NMS, and the cascaded device data is configured on the DHCP server.
2. After the cascaded device is powered on, it automatically initiates a DHCP request to the DHCP server. The DHCP server assigns an IP address to the cascaded device according to the mapping between the IP and MAC addresses of the cascaded device.
3. The cascaded device requests a configuration file from the FTP or Trivial File Transfer Protocol (TFTP) server and loads the configuration file.

NOTE

This implements remote software commissioning for the cascaded device. FTP is recommended, meeting the requirement of higher security.

NAC-based Remote Software Commissioning Using GE Upstream Transmission

The networking of NAC-based remote software commissioning using GE upstream transmission is shown in Figure 21-2 and Figure 21-3.

Figure 21-2 Networking of NAC-based remote software commissioning using GE upstream transmission -1

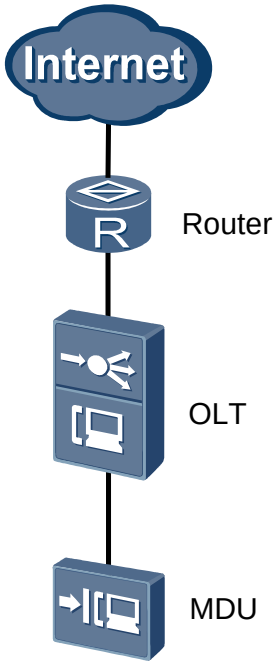
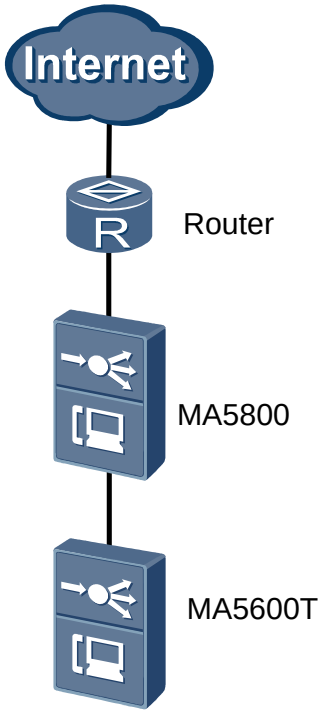


Figure 21-3 Networking of NAC-based remote software commissioning using GE upstream transmission -2



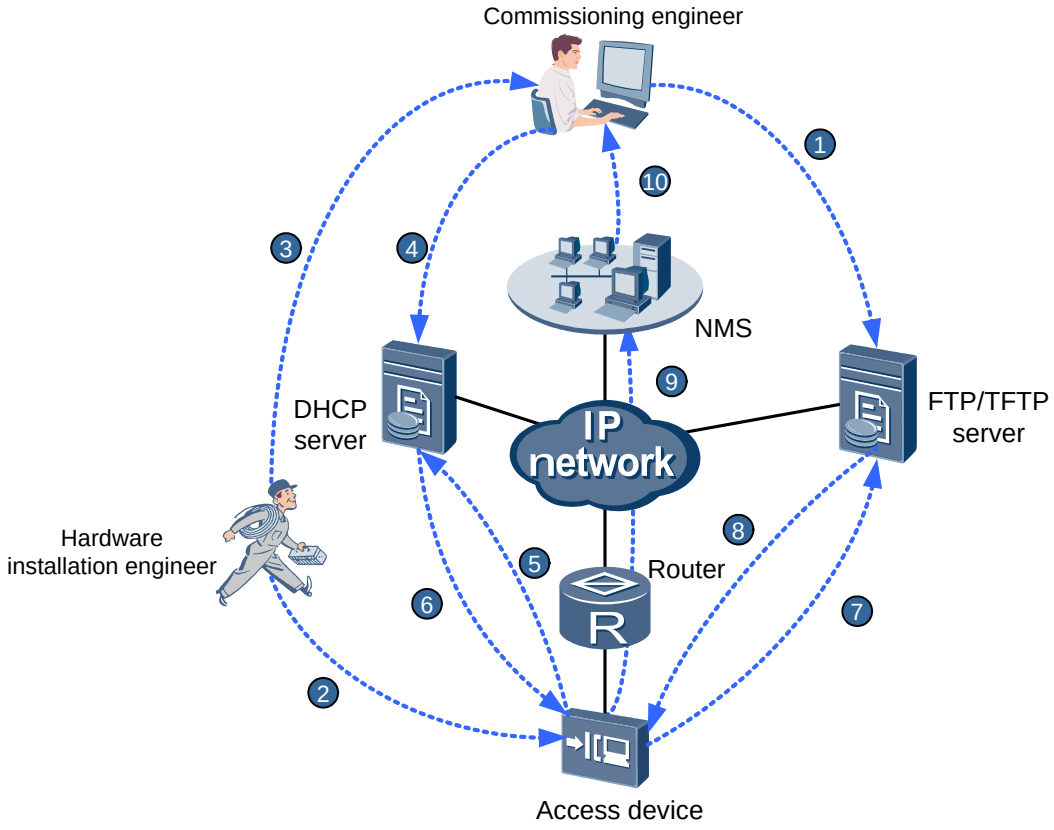
This solution applies to two scenarios: the database of the cascaded device is empty and the database of the cascaded device is not empty.

- If the database of the cascaded device is empty, commissioning engineers can use NAC to remotely configure network management channel parameters and commission the cascaded device. NAC enables the NMS to communicate with the cascaded device through the Internet.
- If the database of the cascaded device is not empty, commissioning engineers can remotely commission the cascaded device after running transparent channel commands to log in to the cascaded device.

21.2 Principles (Based on DHCP)

Figure 21-4 shows the implementation of DHCP-based remote software commissioning using GE upstream transmission.

Figure 21-4 Implementation of DHCP-based remote software commissioning using GE upstream transmission



NOTE

DHCP-based remote software commissioning using GE upstream transmission is configured on the upper-layer devices of access network devices. Therefore, this commissioning type is not described in the remainder of this document.

The process is as follows:

1. The commissioning engineer prepares the configuration file and policy description file and uploads them to the FTP or TFTP server. FTP is recommended, meeting the requirement of higher security.

NOTE

If different devices use different configuration file, the name of the configuration file can be [IP]_cfgfile.txt or [MAC]_cfgfile.txt naming format.

If the devices with same type use same configuration file, the name of the configuration file can be string naming format.

- "IP" is the planned IP address of the MDU. For example, if the IP address is 192.168.10.10, the configuration file of the MDU is named **192.168.10.10_cfgfile.txt**.
- "MAC" is the MAC address of the MDU. For example, if the MAC address is 00-e0-fc-11-ab-ee, the configuration file of the MDU is named **00-e0-fc-11-ab-ee_cfgfile.txt**.
- When the name of the configuration file use the string naming format, for example, if the name of the configuration file is MA5600TscuncfgFile.txt, the devices with this type use this configuration file.

In the [IP]_cfgfile.txt naming format, the IP address of the cascaded device must be unique. To prevent IP address conflict, bind the IP address of the device to its MAC address.

The policy description file must be an .ini file and contain the path and name of the configuration file. The name of the configuration file can be in [IP]_cfgFile.txt, [MAC]_cfgFile.txt, or [string].txt format.

The format and content of a policy description file must be correct. Otherwise, the remote software commissioning flow may be interrupted. If this happens, operators must remotely log in to the cascaded device using the IP address allocated by the DHCP server and reset the device. Alternatively, operators reset the cascaded device onsite. Then, the remote software commissioning flow can be restarted.

One site can use only one policy description file.

The following provides an example of a site policy description file in the [IP]_cfgfile.txt naming format:

```
[MA5600T]                                   //Device type
TargetVersion=MA5600T V800R016C10          //Device version
[H801SCUN]                                  //Control board type
PacketFile="Dir\scun_packetfile.bin"            //IO board package file to be
loaded to the control board
CfgFile="Dir\[IP]_cfgFile.txt"                  //Configuration file to be loaded to
the control board, in the [IP]_cfgfile.txt naming format
```

2. The hardware installation engineer installs the device hardware and powers on the device.

NOTE

Do not remove a board from or insert a board into the cascaded device during device startup.

3. The hardware installation engineer records and reports the MAC address of the cascaded device and site information to the commissioning engineer.
4. The commissioning engineer creates a mapping between the MAC address, management IP address, and physical location of the cascaded device, binds the IP and MAC addresses of the device on the NMS, and configures the device data on the DHCP server.
5. When the LINK indicator on the cascaded device is steady on, the upstream optical path of the cascaded device has been set up. Then, the cascaded device automatically initiates a DHCP request to the DHCP server.
6. The DHCP server assigns an IP address to the cascaded device according to the configuration on the NMS. In addition, the DHCP server writes FTP or TFTP data into DHCP Option packets and sends the packets to the cascaded device. The FTP or TFTP data includes the IP address of the FTP or TFTP server, user account, user password, and policy description file name.

   Table 21-1 lists the data carried in the DHCP Option fields.

Table 21-1 Data carried in the DHCP Option fields

DHCP Option Field	Data

DHCP Option Field	Data
DHCP Option 3	Gateway list
DHCP Option 6	Domain name server (DNS) list
DHCP Option 15	Domain name
DHCP Option 66	TFTP server name
DHCP Option 67	Policy description file name
DHCP Option 141	FTP user name
DHCP Option 142	FTP user password
DHCP Option 143	IP address of the FTP server
DHCP Option 146	Intermediate policy file name description in netfiles
DHCP Option 150	IP address of the TFTP server

7. The cascaded device sets its IP address and gateway and requests a configuration file from the FTP or TFTP server.
8. The FTP or TFTP server issues a configuration file to the cascaded device.
9. The cascaded device automatically loads the configuration file and configures itself with the data. The configuration takes effect after the cascaded device restarts. After restarting, the cascaded device sends a trap to the NMS, informing the NMS that the cascaded device is online.

NOTE

If the loading is failure, the system retry it for 24 hours automatically. If the loading fails after 24 hours all the same, the loading is stopped.

10. The commissioning engineer receives the cascaded device online trap using the NMS.

    The network management channel has been created, and the cascaded device can be remotely managed using the NMS.

21.3 Principles (Based on NAC)

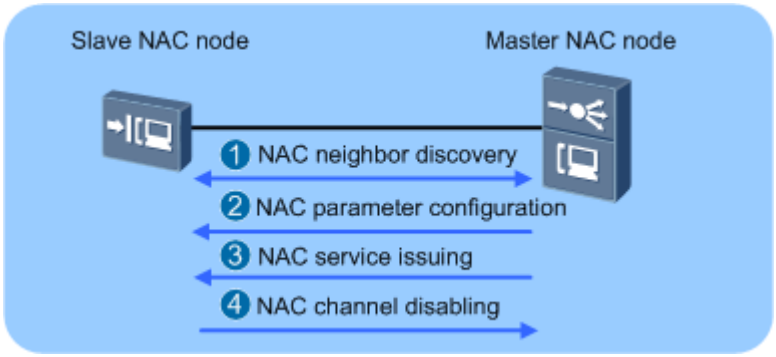
21.3.1 Basic Concepts

Neighbor automatic communication (NAC) is a simple Layer 2 proprietary protocol developed by Huawei. It is used to obtain remote login parameters or basic network management parameters of a newly deployed device from the upper-layer device physically connected to the device.

In NAC, the upper-layer device that physically connects to a newly deployed device functions as a master NAC node, and the newly deployed device functions as a slave NAC node. The NAC-enabled port on the master node is the master NAC port, and the NAC-enabled port on the slave node is the slave NAC port.

Figure 21-5 shows the NAC connection process.

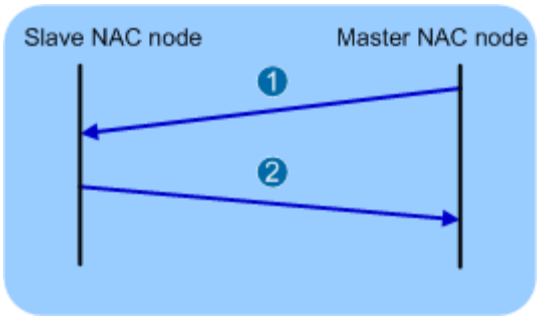
Figure 21-5 NAC connection process



NAC Neighbor Discovery

Figure 21-6 shows the NAC neighbor discovery process.

Figure 21-6 NAC neighbor discovery process



1. Two minutes after the master NAC port status is changed from **down** to **up**, the master NAC node sends a NAC neighbor discovery packet to the slave NAC node.
2. After receiving the discovery packet, the slave NAC node analyzes the packet and sends a response packet to the master NAC node. After the master NAC node verifies that the peer device supports NAC and does not have any service settings, the NAC neighbor discovery process is complete.

NAC Parameter Configuration

NAC parameters include the basic parameters used for remotely logging in to the slave NAC node or for setting up an IP channel for the NMS to manage the slave NAC node. The parameters are described as follows:

- IP channel parameters: include the management VLAN ID, management IP address, and default gateway.
- Network management channel parameters: include the SNMP version, read community name, write community name, IP address of the target host for receiving traps, number of the UDP source port for sending traps, and body name of an SNMP message.

The master NAC node issues NAC parameters to a slave NAC node in offline or online mode.

- Offline mode: Commissioning engineers preconfigure NAC parameters on the master NAC port. After NAC neighbor discovery process is complete, the master NAC node issues the NAC parameters in packets to the slave NAC node.
- Online mode: After the master NAC node discovers a NAC neighbor, commissioning engineers configure NAC parameters on the master NAC port, and the master NAC node issues the NAC parameters in packets to the slave NAC node.

Service Configuration Issuing

After a network management channel for a slave NAC node is set up, commissioning engineers can issue service configurations to the slave NAC node through the CLI or NMS over this channel.

NOTE

If the status of the port receiving remote software commissioning configurations is changed from up to down, the slave NAC node with NAC enabled automatically clears the issued configurations. Therefore, when an exception occurs during configuration issuing, run the **shutdown** and **undo shutdown** commands on the master NAC port. Then, the master NAC node triggers an automatic discovery and configuration again for remote troubleshooting.

NAC Channel Disabling

NOTE

To enhance device security, the slave NAC node disables the NAC channel immediately after commissioning engineers perform non-query operations on the slave NAC node through the CLI or NMS. After the NAC channel is disabled, the issued configurations are saved and the master NAC node cannot change the slave NAC node configurations through NAC.

The slave NAC node automatically disables NAC in any of the following scenarios:

- The slave NAC node has been configured before being deployed onsite.
- The slave NAC node has been commissioned onsite.
- The network management channel has been created through NAC and service configurations have been remotely issued through this channel.

21.3.2 Principles

NAC-based remote software commissioning using GE upstream transmission applies to two scenarios: the database of the cascaded device is empty and the database of the cascaded device is not empty. The principles used in the two scenarios are different.

NOTE

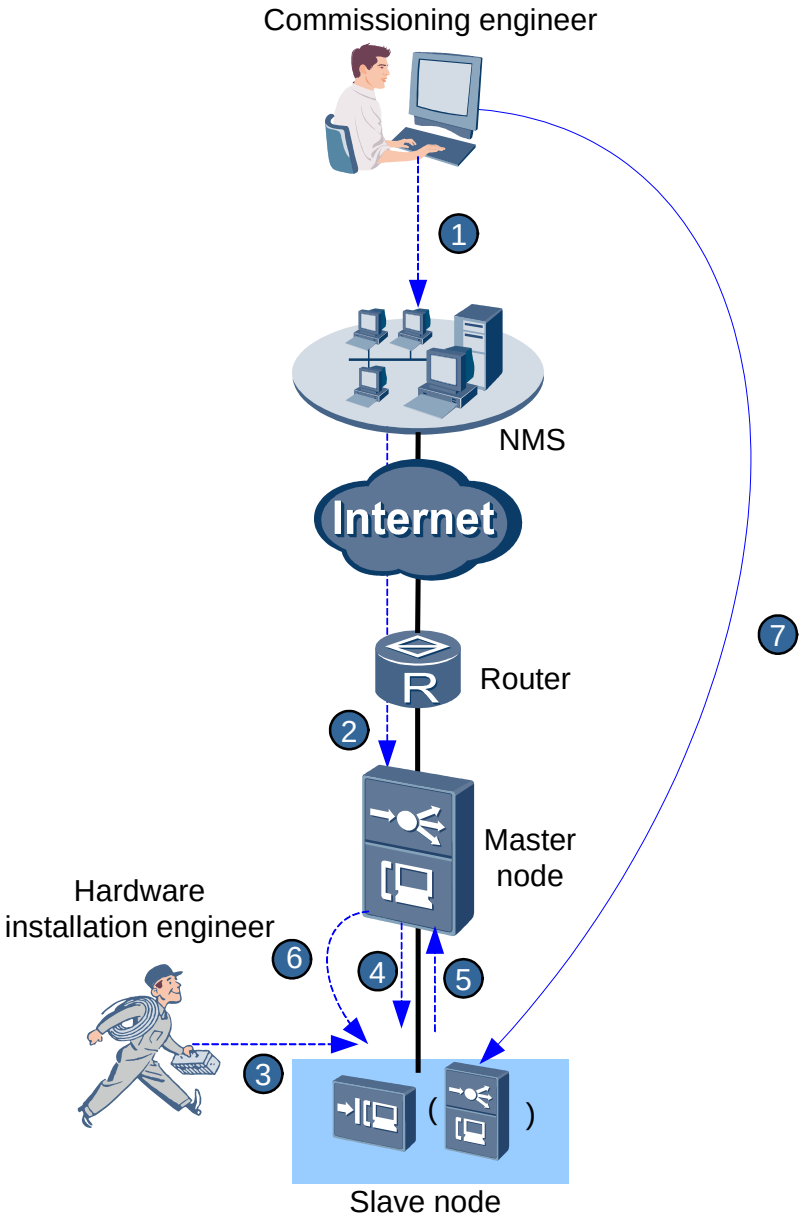
- The cascading device is the master node and the cascaded device is the slave node. NAC must be enabled on both master and slave nodes. To do so, run the **nac enable** command.
- When enabling NAC on the master node, select **security** to enable the security mode. In this case, the master NAC node encrypts remote software commissioning parameters before sending them to the slave NAC node. This improves parameter security.
- The slave node does not save the management parameters issued through NAC. To save the parameters, run the **save** command. After saving the parameters, the NAC function will be off automatically.

Principles Used If the Slave Node Database Is Empty

This scenario involves two sub-scenarios: with NMS and without NMS.

Figure 21-7 shows the principles of NAC-based remote software commissioning using GE upstream transmission with NMS.

Figure 21-7 Principles of NAC-based remote software commissioning using GE upstream transmission with NMS



The process is as follows:

1. The commissioning engineer imports the planned data into the NMS. The planned data includes network topology connections, including master node ports, and slave node device management parameters, including the management VLAN, SNMP parameter profile, and NAC parameters.
2. The NMS automatically issues the network management channel data for the slave node device to the master node device.
3. The hardware installation engineer obtains the slave node device from the warehouse and delivers it to the site. Then, the hardware installation engineer installs the slave node device, connects wires for the slave node device, and powers on the slave node device.

NOTE

Do not remove a board from or insert a board into the slave node device during the slave node device startup.

4. After the slave node device is powered on, the status of the master node port connected to the slave node device is changed from **down** to **up**. Then, the master node device sends a NAC neighbor discovery packet to the slave node device.
5. After receiving the discovery packet, the slave node device analyzes the packet and sends a response packet to the master node device.
6. After the master node device verifies that the slave node device supports NAC and does not have any service settings, the NAC neighbor discovery process is complete. Then, the master node device issues NAC parameters and network management channel parameters to the slave node device. After obtaining the network management channel parameters, the slave node device can be remotely logged in to.

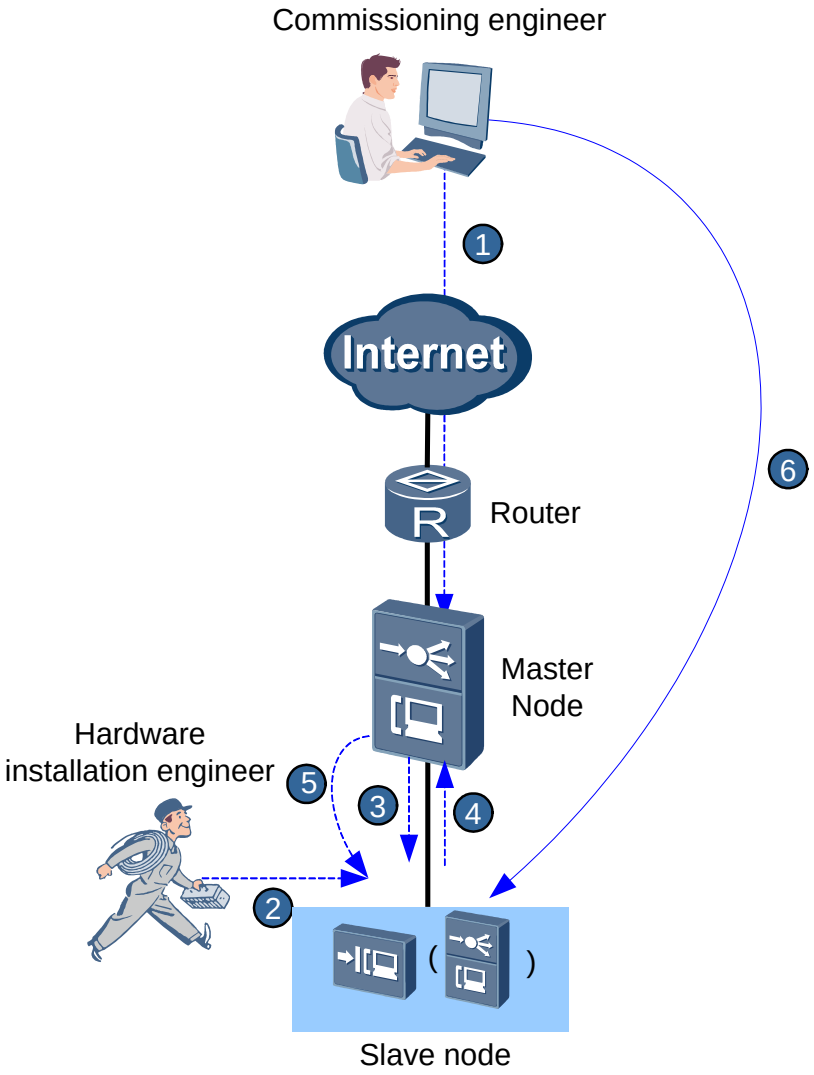
NOTE

Run the **nac disable** command to disable the NAC function of the master node after the slave node obtains the device management parameters. This prevents the master node from issuing the originally planned data to a substituted slave node device.

7. The commissioning engineer issues service configurations to the slave node device through the CLI or NMS.

Figure 21-8 shows the principles of NAC-based remote software commissioning using GE upstream transmission without NMS.

Figure 21-8 Principles of NAC-based remote software commissioning using GE upstream transmission without NMS



The process is as follows:

1. The commissioning engineer logs in to the master node device and configure slave node device management parameters on the master node port connected to the slave node device. The slave node device management parameters include the management VLAN, and NAC parameters.
2. The hardware installation engineer obtains the slave node device from the warehouse and delivers it to the site. Then, the hardware installation engineer installs the slave node device, connects wires for the slave node device, and powers on the slave node device.

NOTE
Do not remove a board from or insert a board into the slave node device during the slave node device startup.

3. After the slave node device is powered on, the status of the master node port connected to the slave node device is changed from **down** to **up**. Then, the master node device sends a NAC neighbor discovery packet to the slave node device.

4. After receiving the discovery packet, the slave node device analyzes the packet and sends a response packet to the master node device.
5. After the master node device verifies that the slave node device supports NAC and does not have any service settings, the NAC neighbor discovery process is complete. Then, the master node device issues NAC parameters to the slave node device. After obtaining the NAC parameters, the slave node device can be remotely logged in to.

NOTE

Run the **nac disable** command to disable the NAC function of the master node after the slave node obtains the device management parameters. This prevents the master node from issuing the originally planned data to a substituted slave node device.

6. The commissioning engineer issues service configurations to the slave node device through the CLI.

Principles Used If the Slave Node Database Is Not Empty

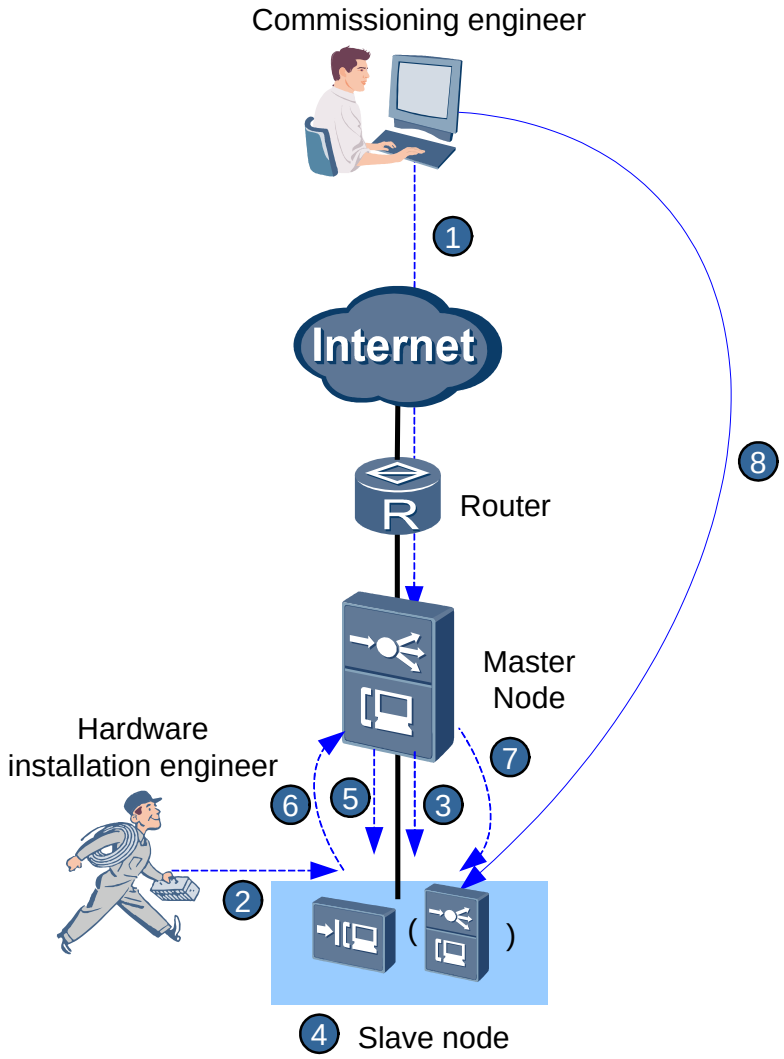
If the slave node database is not empty, commissioning engineers must log in to the slave node device by running transparent channel commands and then remotely commission the slave node device. This scenario involves two sub-scenarios: the slave node database needs to be erased and the slave node database data needs to be retained.

NOTICE

The packets transmitted over the transparent channel are not encrypted and therefore these operations are risky. Exercise caution when performing any operations involving the transparent channel.

Figure 21-9 shows the principles of NAC-based remote software commissioning using GE upstream transmission if the slave node database is not empty and the database needs to be erased.

Figure 21-9 Principles of NAC-based remote software commissioning using GE upstream transmission if the slave node database is not empty and the database needs to be erased



The process is as follows:

1. The commissioning engineer logs in to the master node device and configure slave node device management parameters on the master node port connected to the slave node device. The slave node device management parameters include the management VLAN, and NAC parameters.
2. The hardware installation engineer obtains the slave node device from the warehouse and delivers it to the site. Then, the hardware installation engineer installs the slave node device, connects wires for the slave node device, and powers on the slave node device.

NOTE

Do not remove a board from or insert a board into the slave node device during the slave node device startup.

3. After the slave node device is powered on, the NAC neighbor discovery fails because the slave node database is not empty. Then, the commissioning engineer runs the **transparent on** command to enable the transparent channel from the master node port connected to the slave node device to the slave node device and logs in to the slave node device.

NOTE

- The packet transmission over the transparent channel is based on NAC.
- The transparent channel is enabled on the slave node device by default. To ensure system security or if the transparent channel is not used, run the **transparent remote disable** command to disable it.

4. The commissioning engineer runs the **erase flash data** command to clear the slave node database and runs the **reboot system** command to reset the slave node device.
5. After the slave node device is powered on, the status of the master node port connected to the slave node device is changed from **down** to **up**. Then, the master node device sends a NAC neighbor discovery packet to the slave node device.
6. After receiving the discovery packet, the slave node device analyzes the packet and sends a response packet to the master node device.
7. After the master node device verifies that the slave node device supports NAC and does not have any service settings, the NAC neighbor discovery process is complete. Then, the master node issues NAC parameters to the slave node. After obtaining the NAC parameters, the slave node device can be remotely logged in to.

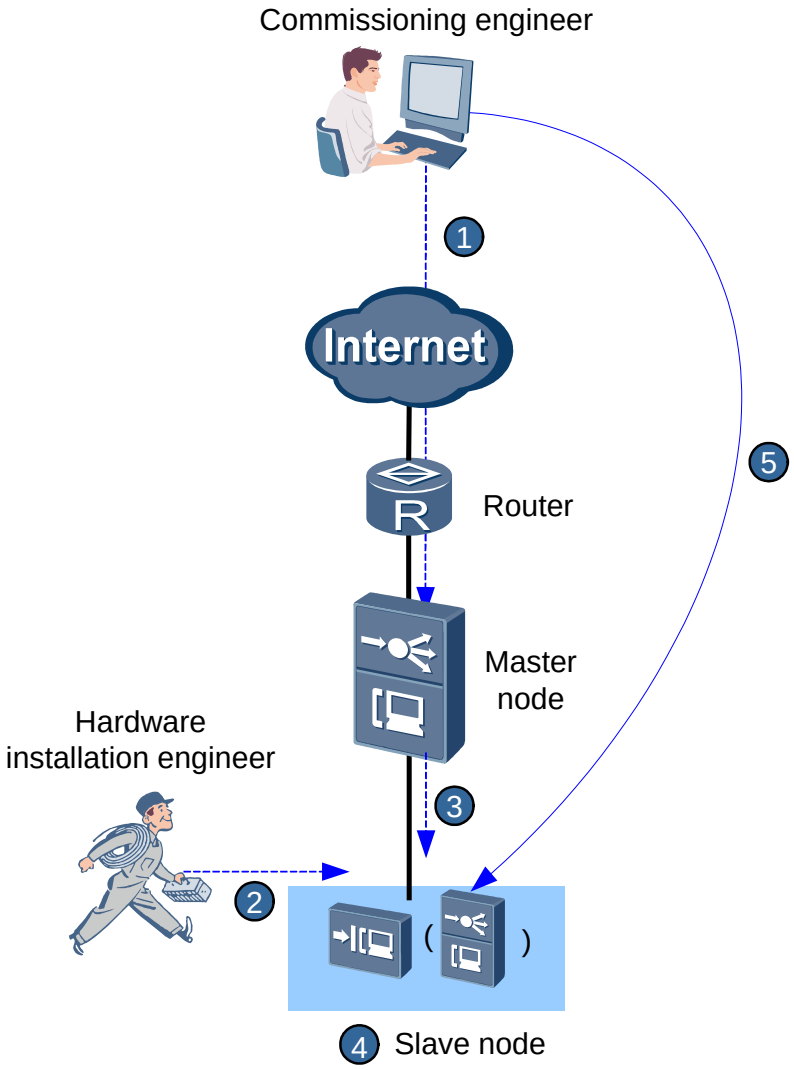
NOTE

- Run the **nac disable** command to disable the NAC function of the master node after the slave node obtains the device management parameters. This prevents the master node from issuing the originally planned data to a substituted slave node device.
- The transparent channel commands for logging in to a slave node are diagnostic commands. After the slave node communicates with the master node at Layer 3, log in to and configure the slave node in remote mode. This prevents exceptions caused by the transparent channel commands from occurring on the slave node.

8. The commissioning engineer issues service configurations to the slave node device through the CLI.

Figure 21-10 shows the principles of NAC-based remote software commissioning using GE upstream transmission if the slave node database is not empty and the database data needs to be retained.

Figure 21-10 Principles of NAC-based remote software commissioning using GE upstream transmission if the slave node database is not empty and the database data needs to be retained



The process is as follows:

1. The commissioning engineer logs in to the master node device.
2. The hardware installation engineer obtains the slave node device from the warehouse and delivers it to the site. Then, the hardware installation engineer installs the slave node device, connects wires for the slave node device, and powers on the slave node device.

NOTE

Do not remove a board from or insert a board into the slave node device during the slave node device startup.

3. After the slave node device is powered on, the NAC neighbor discovery fails because the slave node database is not empty. Then, the commissioning engineer runs the **transparent on** command to enable the transparent channel from the master node port connected to the slave node device to the slave node device and logs in to the slave node device.

NOTE

- The packet transmission over the transparent channel is based on NAC.

- The transparent channel is enabled on the slave node device by default. To ensure system security or if the transparent channel is not used, run the **transparent remote disable** command to disable it.

4. The commissioning engineer configures slave node device management parameters on the slave node device. The slave node device management parameters include the management VLAN, ip address, and gateway address. After the configuration, the slave node device can be remotely logged in to.

NOTE

- Run the **nac disable** command to disable the NAC function of the master node after the slave node obtains the device management parameters. This prevents the master node from issuing the originally planned data to a substituted slave node device.
- The transparent channel commands for logging in to a slave node are diagnostic commands. After the slave node communicates with the master node at Layer 3, log in to and configure the slave node in remote mode. This prevents exceptions caused by the transparent channel commands from occurring on the slave node.

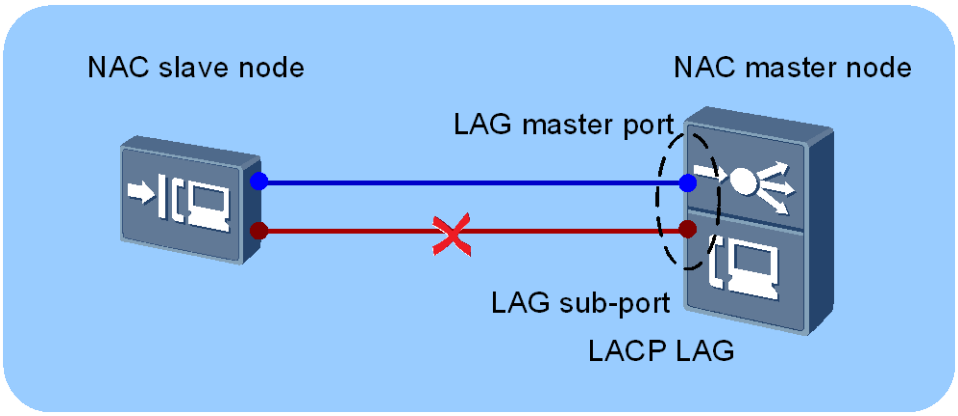
5. The commissioning engineer issues service configurations to the slave node device through the CLI.

21.3.3 LAG Application on a NAC Master Node

LACP LAG

A NAC slave node automatically creates a link aggregation group (LAG) and runs the Line Aggregation Control Protocol (LACP) only for the neighbor automatic communication (NAC) slave port that is connected to the master port in the LAG. The NAC slave port connected to the sub-port in a LAG will be blocked due to failed LACP negotiation, as shown in Figure 21-11.

Figure 21-11 NAC application in a LACP LAG



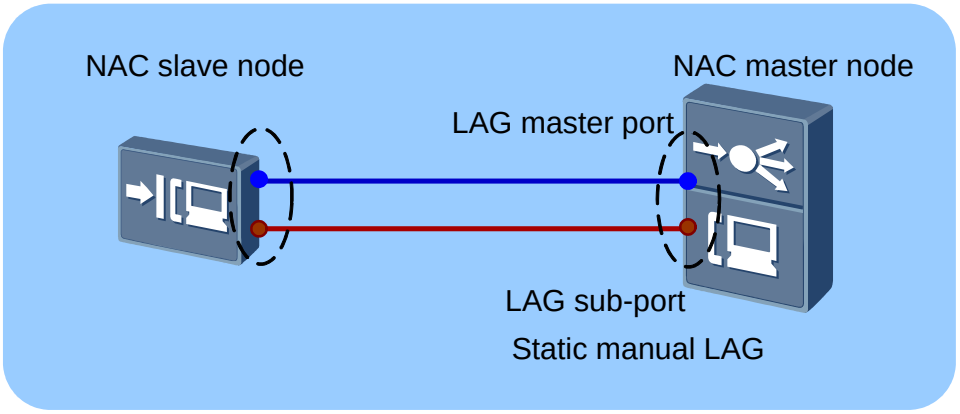
Static Manual LAG

NAC can be configured only on the master port in a LAG.

NAC automatically creates a LAG on the NAC slave node and adds the port connected to the LAG of the NAC master node to the LAG after the port is detected. The added Ethernet ports may be different from those in the network plan because the port addition is affected by the port connection sequence and the time at which the NAC slave node disables the NAC channel. Therefore, manually adjust the added ports.

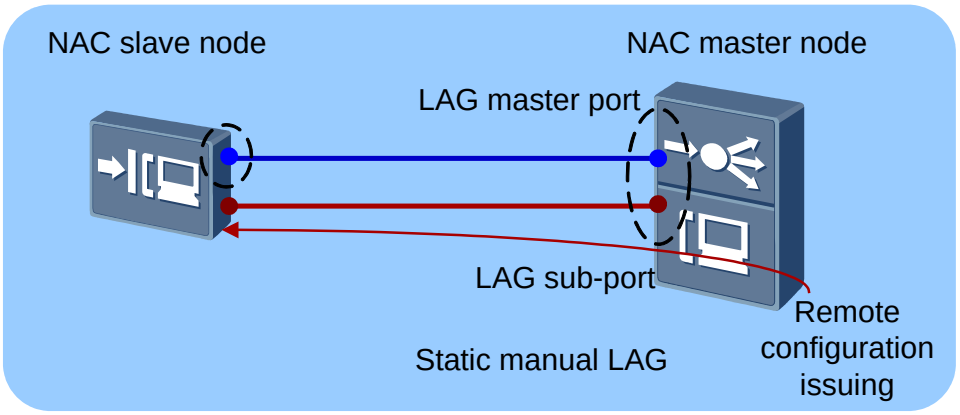
Figure 21-12 shows an overall process of how NAC is applied on a static manual LAG. The NAC master node issues a remote commissioning message that contains port information about the LAG. After receiving the message, the NAC slave node creates a LAG on the Ethernet port that receives the NAC configuration. In addition, the NAC slave node checks the master port information received by all the slave ports (these ports receive the information after startup). If the sub-port ID in the master port information received by a slave port is the same as that in the remote commissioning message, the NAC slave node automatically adds this slave port to the LAG of the slave node.

Figure 21-12 NAC application on a static manual LAG-1



As shown in Figure 21-13, if a port in the LAG is activated after the NAC slave node disables the NAC channel, the NAC slave node does not automatically add the port to the LAG. When forwarding downstream streams, if the NAC master node selects the link connecting to this port, the streams are discarded on the NAC slave node. As a result, the remote commissioning is interrupted. To resolve this issue, run the **shutdown** command on the NAC slave node to disable sub-ports in the LAG. After the remote commissioning channel recovers, manually add all required ports to the static LAG on the NAC slave node and run the **undo shutdown** command to enable the sub-ports in the LAG.

Figure 21-13 NAC application on a static manual LAG-2



21.4 Configuring NAC-based Remote Software Commissioning Using GE Upstream Transmission

The OLT supports NAC, which allows an MDU cascaded to the OLT using a GE port and an MA560xT cascaded to an MA5800 using a GE port to automatically obtain network management parameters. Therefore, commissioning engineers do not need to visit a site for MDU commissioning.

Prerequisites

Management parameters, including the inband management VLAN ID, management IP address, and SNMP profile parameters, have been configured on the cascading device for cascaded devices. (SNMP profile parameters are required if the cascaded devices must be managed through the NMS.)

Context

This section uses the scenario where the OLT is a deployed device and the MDU is a newly deployed device as an example to describe the configuration process. For the scenario where an MA560xT is cascaded to an MA5800, the configuration process is the same as that used in scenario where an MDU is cascaded to an OLT, excepting that the former scenario does not support automatic deployment policy files. Therefore, the configuration process of the latter scenario will not described in the remainder of the document.

NOTE
The MDU does not save the management parameters issued through NAC. To save the parameters, run the **save** command. After saving the parameters, the NAC function will be off automatically.

Procedure

- The MDU database is empty.
  a. Enable NAC on the OLT.

     The OLT port connected to the MDU is 0/3/0.

     ```
     huawei(config)#nac enable master 0/3 0
     ```

NOTE
When enabling NAC on the OLT, select **security** to enable the security mode. In this case, the master NAC node encrypts remote software commissioning parameters before sending them to the slave NAC node. This improves parameter security.

  b. Configure MDU device management parameters on the OLT.

     The inband management VLAN ID is 8, the management IP address is 192.168.50.2, the IP address of the gateway is 192.168.50.254, and the SNMP profile ID is 64. These parameters can be configured using NAC or RN commands.

NOTE
MDU device management parameters cannot be configured using both NAC and RN commands. Otherwise, the OLT will generate an error message.

     Configure MDU device management parameters using NAC commands.

     ```
     huawei(config)#nac config master 0/3/0 slave vlan 8 priority 3 ip-address
     192.168.50.2 mask 255.255.255.0 gateway 192.168.50.254 snmp-profile-id 64
     ```

     Configure MDU device management parameters using RN commands.

     ```
     huawei(config)#rn ipconfig 0/3/0 ip-address 192.168.50.2 mask 255.255.255.0
     gateway 192.168.50.254 vlan 8 priority 3
     ```

```
huawei(config)#rn snmp-config all snmp-index 0 profile-id 64
```

c. Check whether the MDU device management parameters have been configured on the OLT.

Query parameter configurations using NAC or RN commands.

Query parameter configurations using NAC commands.

```
huawei(config)#display nac configuration master
{ <cr>|frameid/slotid/portid<S><Length 5-18>|frameid/slotid<S><Length
3-15> }:

  Command:
         display nac configuration master

------------------------------------------------------------------------------
---
  Frame ID/Slot ID/Port ID :  0/3/0
  VLAN                  :  8
  VLAN priority          :  3
  IP address             :  192.168.50.2
  Mask                  :  255.255.255.0
  Gateway               :  192.168.50.254
  Snmp-profile-id        :  64
  Snmp-profile-name      :  snmp-profile_64

------------------------------------------------------------------------------
---
  Total: 1
```

Query parameter configurations using RN commands.

```
huawei(config)#display rn info 0/3/0
  Command:
       display rn info 0/3/0

------------------------------------------------------------------------------
----
  F/S/P            :  0/3/0
  Equipment ID      :  XXXXXX  //"XXXXXX" is the MDU name.
  Uplink type       :  ETH
  MAC/SN           :  485754430CA7AB05
  Run state         :  Offline
  Config state      :  -
  IP               :  192.168.50.2
  Subnet mask       :  255.255.255.0
  Gateway          :  192.168.50.254
  Manage VLAN       :  8
  Manage priority    :  3
  Deploy profile ID  :  -
  Deploy profile name :  -
  ----------------------------------------------------------------
  Index  SNMP       SNMP                         Effective
        Profile ID  Profile name                   Flag
  ----------------------------------------------------------------
  0     64        snmp-profile_64               No
  ----------------------------------------------------------------
```

```
------------------------------------------------------------------------
----
```

d. Configure an automatic deployment policy profile.

Perform this step if an MDU is cascaded to an OLT. Profile configurations are as follows:

- Profile name: **deploy-profile_1**
- MDU configuration file name: **cfgfile.txt**
- MDU IP address: 10.10.10.10
- File transfer protocol: FTP
- User name: **user1**
- Password: **user123** (not displayed when the password is entered)

```
huawei(config)#rn deploy-profile add  profile-id 1 profile-name
deploy-profile_1 filename cfgfile.txt ip 10.10.10.10 ftp user user1
```

e. Bind the automatic deployment policy profile to the MDU.

Perform this step if an MDU is cascaded to an OLT.

Bind the profile based on ports:

```
huawei(config)#rn deploy-config 0/3/0 profile-id 1
```

Bind the profile based on device types:

```
huawei(config)#rn deploy-policy MA56XX profile-id 1 //Change "MA56XX" to the
actual MDU type.
```

NOTE

The ports-based profile binding takes precedence over the device types-based profile binding.

f. After the MDU is powered on, it exchanges data with the OLT to obtain the MDU device management parameters and service configuration parameters. Wait for 2 minutes. Then, run the **telnet** command to remotely log in to the MDU.

g. Query the device connected to the OLT cascading port.

The device can be queried using NAC or RN commands.

Query the device using NAC commands.

```
huawei(config)#display nac slave info detail

------------------------------------------------------------------------
---
  NAC master port         :  0/3/0
  NAC auto-find fsm state   :  Established
  NAC auto-find result     :  Found
  Multiple slave node exist :  No
  Slave node device type    :  HUAWEI XXXXXX  //"XXXXXX" is the MDU name.
  Slave node mac-address    :  00-11-01-01-00-52
  NAC slave port          :  0/0/1
  NAC config fsm state     :  INIT
  NAC config result       :  Added successfully
  NAC config fail reason    :  -

------------------------------------------------------------------------
```

```
---
```

Query the device using RN commands.

```
huawei(config)#display rn info 0/3/0
  Command:
       display rn info 0/3/0


------------------------------------------------------------------------------
----
  F/S/P             :  0/3/0
  Equipment ID       :  XXXXXX  //"XXXXXX" is the MDU name.
  Uplink type        :  ETH
  MAC/SN            :  485754430CA7AB05
  Run state          :  Online
  Config state       :  Normal
  IP                :  192.168.50.2
  Subnet mask        :  255.255.255.0
  Gateway           :  192.168.50.254
  Manage VLAN        :  8
  Manage priority    :  3
  Deploy profile ID  :  1
  Deploy profile name :  deploy-profile_1
  ----------------------------------------------------------------
  Index  SNMP       SNMP                          Effective
        Profile ID  Profile name                    Flag
  ----------------------------------------------------------------
  0     64        snmp-profile_64                 No
  ----------------------------------------------------------------


------------------------------------------------------------------------------
----
```

h. (Optional) Disable NAC on the OLT.

Perform this operation after the MDU obtains the device management parameters. This prevents the OLT from issuing the originally planned data to a substituted MDU.

```
huawei(config)#nac disable master 0/3 0
```

- The MDU database is not empty.

NOTICE

The packets transmitted over the transparent channel are not encrypted and therefore these operations are risky. Exercise caution when performing any operations involving the transparent channel.

The following section only describes the configuration process if the MDU database must be erased. The configuration is simple if the MDU database does not need to be erased because you only need to perform the operations in b to log in to the MDU and configure device management parameters on the MDU. Therefore, this configuration process is not described in the remainder of this document.

NOTE

The transparent channel is enabled on the MDU by default. To ensure system security or if the transparent channel is not used, run the **transparent remote disable** command to disable it.

a. Perform steps a through e.

b. After the MDU is powered on, enable the transparent channel on the OLT port, and log in to the MDU through the transparent channel.

The OLT port connected to the MDU is 0/3/0.

```
huawei(config)#diagnose
huawei(diagnose)%%transparent on 0/3/0
>>User name:huawei  //Enter the user name of the MDU.
>>User password:   //Enter the user password.


  Huawei Integrated Access Software (huawei).
  Copyright(C) Huawei Technologies Co., Ltd. 2002-2013. All rights reserved.


------------------------------------------------------------------------
---
  User last login information:

------------------------------------------------------------------------
---
  Access Type : Serial
  IP-Address  : --
  Login  Time : 2013-07-21 01:04:17+08:00
  Logout Time : 2013-07-21 05:19:35+08:00

------------------------------------------------------------------------
---

  User root has used a default password. Change the password in time.
```

c. Erase the MDU database.

```
huawei>enable
huawei#erase flash data
  This command will clear the active board data that has been loaded or saved

  Please remember to backup the system configuration data
  Are you sure to continue? (y/n)[n]:y
  Command executes successfully, and will take effect after active board
rebooted
```

d. Reset the MDU.

```
huawei#reboot system
  Please check whether data has saved, the unsaved data will lose if reboot
system, are you sure to reboot system? (y/n)[n]:y
```

e. After the MDU is reset, OLT commands are displayed. Wait for 2 minutes, then check the configuration.

The remote software commissioning configuration can be queried using NAC or RN commands.

Query the remote software commissioning configuration can be queried using NAC commands.

```
huawei(diagnose)%%quit
huawei#display nac slave info detail

------------------------------------------------------------------------------
---
  NAC master port         :  0/3/0
  NAC auto-find fsm state   :  Established
  NAC auto-find result     :  Found
  Multiple slave node exist :  No
  Slave node device type    :  HUAWEI XXXXXX  //"XXXXXX" is the MDU name.
  Slave node mac-address    :  00-11-01-01-00-52
  NAC slave port          :  0/0/1
  NAC config fsm state     :  INIT
  NAC config result       :  Added successfully
  NAC config fail reason    :  -

------------------------------------------------------------------------------
---
  Total: 1
```

Query the remote software commissioning configuration can be queried using RN commands.

```
huawei(config)#display rn info 0/3/0
  Command:
       display rn info 0/3/0

------------------------------------------------------------------------------
----
  F/S/P             :  0/3/0
  Equipment ID       :  XXXXXX  //"XXXXXX" is the MDU name.
  Uplink type        :  ETH
  MAC/SN            :  485754430CA7AB05
  Run state          :  Online
  Config state       :  Normal
  IP                :  192.168.50.2
  Subnet mask        :  255.255.255.0
  Gateway           :  192.168.50.254
  Manage VLAN        :  8
  Manage priority    :  3
  Deploy profile ID   :  1
  Deploy profile name :  deploy-profile_1
  ---------------------------------------------------------------
  Index  SNMP       SNMP                          Effective
         Profile ID  Profile name                   Flag
  ---------------------------------------------------------------
  0     64         snmp-profile_64                No
  ---------------------------------------------------------------

------------------------------------------------------------------------------
----
```

f. Run the **telnet** command to remotely log in to the MDU.

NOTE

The transparent channel commands for logging in to an MDU are commissioning commands. After the MDU communicates with the OLT at Layer 3, log in to and configure the MDU in remote mode. This prevents exceptions caused by the transparent channel commands from occurring on the MDU.

----End

21.5 Reference Standards and Protocols

The following table lists the reference standards and protocols of remote software commissioning using GE upstream transmission.

Standard	Description
IETF RFC 2131	Dynamic Host Configuration Protocol
IETF RFC 1533	DHCP Options and BOOTP Vendor Extensions
IEEE 802.3ah-2004	IEEE Standard for Information technology- Telecommunications and information exchange between systems- Local and metropolitan area networks- Specific requirements Part 3: Carrier Sense Multiple Access with Collision Detection (CSMA/CD) Access Method and Physical Layer Specifications Amendment: Media Access Control Parameters, Physical Layers, and Management Parameters for Subscriber Access Networks

22 Centralized Management for GE Remote Extended Subracks in FTTB or FTTC Scenarios

About This Chapter

With the development of high-bandwidth copper line technologies at the user end, such as VDSL2 and Vectoring, reconstruction of the last mile on copper line networks suffices to provide services at rates near 100 Mbit/s. Under such a background, more and more small-capacity devices are deployed at the remote end. If these devices are still managed as standalone NEs, the management scale is excessively large and the operating expense (OPEX) is high. The MA5600T/MA5603T provides the centralized management for GE remote extended subracks solution to efficiently manage devices, and reduce OPEX and total cost of operation (TCO).

22.1 Introduction to Centric Management for GE Remote Extended Subrack

Centralized management for GE remote extended subracks in FTTB or FTTC scenarios: The digital subscriber line access multiplexer (DSLAM) or multi-service access node (MSAN) deployed in the central office (CO) functions as the master subrack and connects to remote small-capacity devices (remote extended subracks, such as the MA5623AR) using GE extending boards. The xDSL service board of a remote extended subrack works with the xDSL service board of the master subrack to provide xDSL services within the coverage reach. In this manner, remote extended subracks are no longer standalone NEs (no longer allocated independent management IP addresses), but are managed by the master subrack. These remote extended subracks are regarded as remote service boards of the master subrack and have the same functions and features as those of the master subrack.

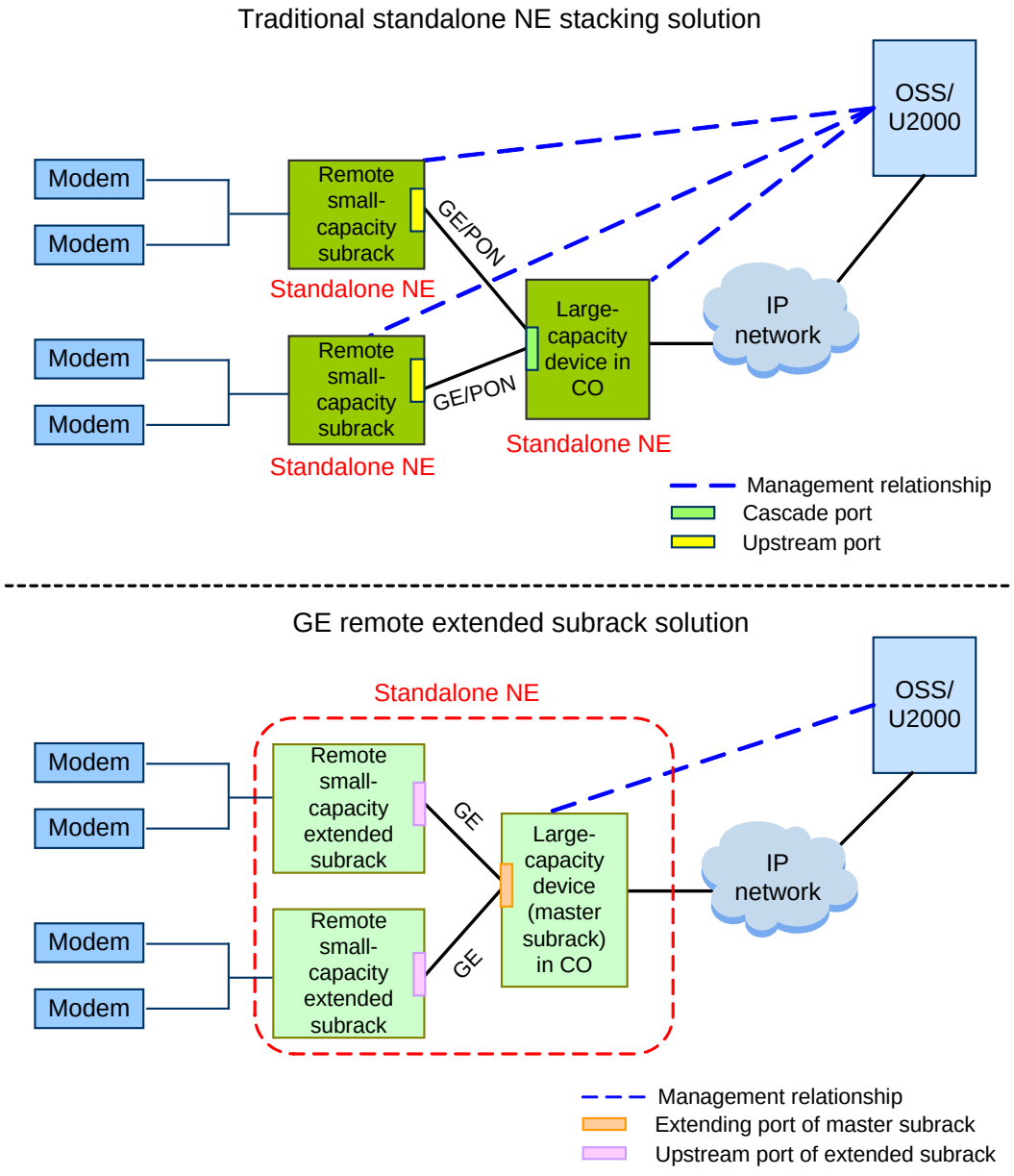
Advantages of the centralized management for GE remote extended subracks:

- The user coverage of the system expands, but management objects do not increase.
- Software commissioning is not required for remote extended subracks and one onsite operation suffices.
- Service provisioning interfaces of remote extended subracks remain the same. The user interface for extended subracks is the same as that for a service board of the master

subrack, enabling consistent user experience. Each small-capacity device deployed at the remote end can be regarded as a new service board deployed on the master subrack. The small-capacity device does not need to interconnect with the upper-layer OSS or NMS system, reducing operating expense (OPEX) and total cost of operation (TCO) for carriers.

Figure 22-1 shows the differences between the centralized management for GE remote extended subracks solution and traditional standalone NE stacking solution.

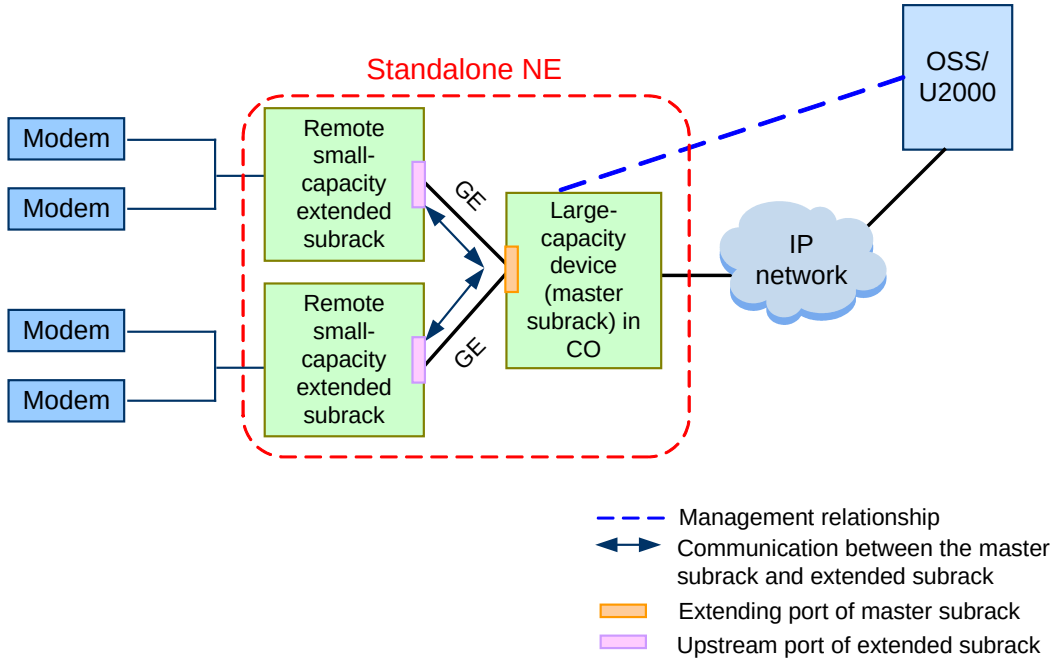
Figure 22-1 Comparison between the two solutions



22.2 GE remote extended subrack Management

According to the centralized management for GE remote extended subracks solution, the master subrack (MA5600T/MA5603T) and the extended subracks are regarded as one NE by the OSS and NMS. Therefore, only one management IP address is planned and the service ports on the extended subrack are managed and maintained through the master subrack.

Figure 22-2 Centralized management for GE remote extended subracks



- As shown in Figure 22-2, the service boards on the remote extended subrack are regarded as remote boards of the master subrack, because they can exchange messages with the master subrack, support the same user experience as the VDSL2 service boards on the master subrack, and are processed like service boards on the master subrack.
- On the NMS, user ports are still identified by frame ID/slot ID/port ID. Except the frame ID is extended for the remote extended subrack, other IDs remain the same.
- Like the master subrack, the remote extended subrack also supports the offline pre-deployment and automatic discovery scenarios.
  - Offline pre-deployment scenario: Run the **network-role** command on the master subrack to set the ETHB extending board to the extend working mode. Then, run the **frame add** command to add the remote extended subrack offline. After being installed and powered on, the extended subrack automatically registers with the master subrack and enters the normal state.

NOTICE

- Before setting the working mode of the ETHB board to **extend**, you should confirm that the auto-negotiation function of the extending port is enabled.
- Switching the working mode of the ETHB board between **extend** and another working mode causes an ETHB board to reset.

  - Automatic discovery scenario: Run the **display network-role** command on the master subrack to check whether the ETHB extending board is working in the extend mode. If not, run the **network-role** command to set the ETHB board to the extend working mode. Then, run the **frame confirm** command to confirm the extended subrack. In this manner, the extended subrack enters the normal state.

NOTE

- The software creates one virtual control board and one virtual service board for the MA5623AR extended subrack. These two boards can be queried by running the **display board** command on the MA5600T/MA5603T. The control board on the master subrack directly manages the two virtual boards, and the software and patch upgrades of all boards on the extended subrack.
- The control board on the extended subrack manages upstream ports and monitors the subrack environment but does not manage service boards.

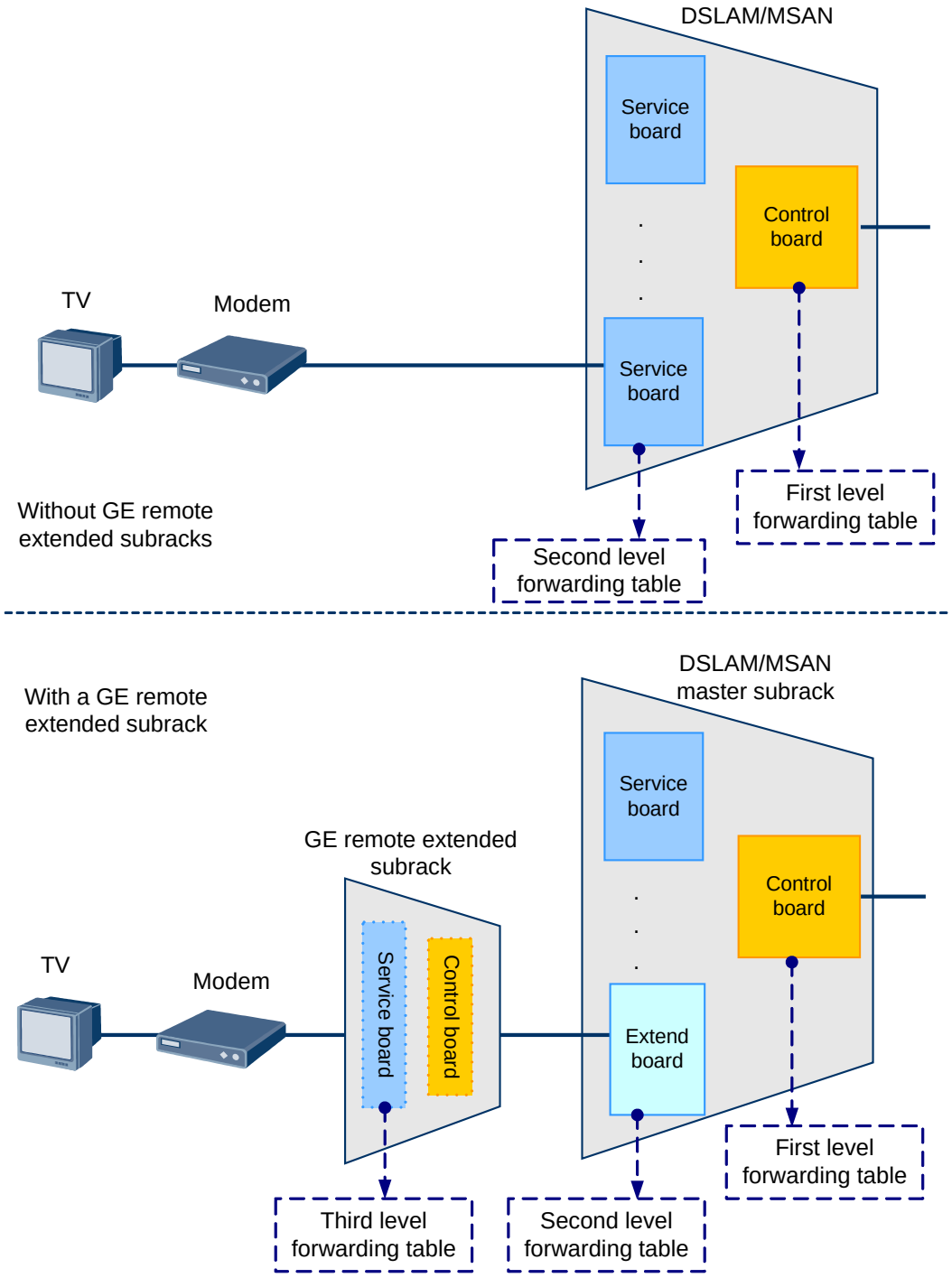
22.3 Working Principles of the GE remote extended subrack

In the GE remote extended subrack implementation, all features are implemented in the same way as those of the MA5600T/MA5603T, except the distributed multicast and Ethernet uplink aggregation features. This topic will detail the working principles of these two features of the GE remote extended subrack.

Working Principles of the Distributed Multicast Feature

As shown in Figure 22-3, after the GE remote extended subrack is added for the MA5600T/MA5603T, the "service board-control board" two-level multicast is extended to "service board on extended subrack-extending board on master subrack-control board on master subrack" three-level multicast. Packets in the three-level multicast are processed in the same way as packets in the two-level multicast. The implementation principle is as follows: The control and extend boards on the master subrack and the service board on the extended subrack create multicast forwarding tables separately. Then, based on the forwarding tables, multicast packets are forwarded from the control board of the master subrack to the extending board, then to the service board on the extended subrack, and finally to users.

Figure 22-3 Implementation of distributed multicast with the extended subrack



During the implementation, except the extended number of multicast users and frame IDs, configuration parameters and multicast feature specifications remain the same.

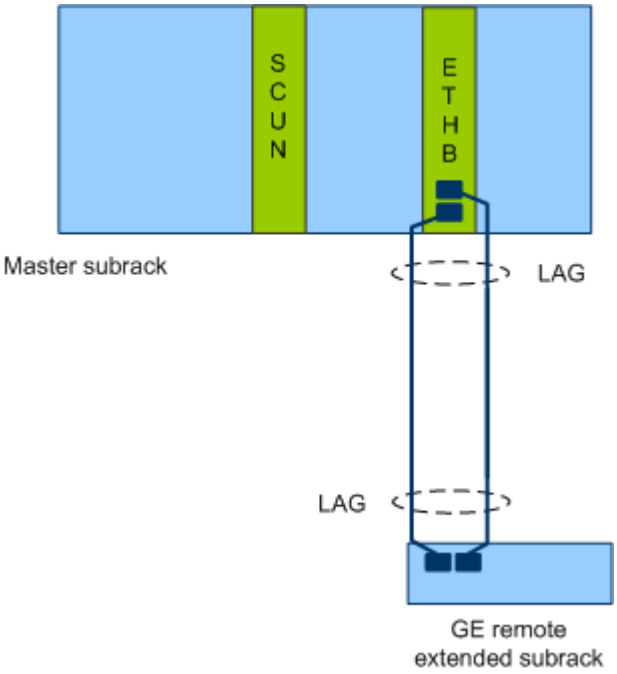
Working Principles of the Ethernet Uplink Aggregation Feature

The link aggregation principles of the GE remote extended subrack are similar to those of the master subrack. For details, see 19.3 Ethernet Link Aggregation.

In the GE remote extended subrack scenario, note the following points:

- Currently, the ETHB extending board supports only intra-board aggregation network and manual aggregation workmode.
- For the uplink aggregation configuration on the master subrack, the ETHB extending port, which is specified during the offline pre-deployment of the extended subrack or where the extended subrack is automatically discovered, must be configured as the master port of a link aggregation group (LAG). However, the extended subrack has no requirements on the master port of the LAG.
- Before adding or deleting a port in a LAG which interconnects with a peer device, run the **shutdown** command to deactivate the slave port of the LAG or remove the fiber from the slave port to prevent a loopback. Note that ports on the ETHB board support the **shutdown** command, but upstream ports on the extended subrack do not. After the LAG is configured, re-activate the slave port or reconnect the fiber to the slave port.

Figure 22-4 Implementation of Ethernet uplink aggregation with the remote extended subrack



22.4 Adding GE Remote Extended Subracks

This topic describes how to add a GE remote extended subrack to the master subrack so that the master subrack can manage the extended subrack.

Data Plan

The following table lists the data plan for interconnection between the master subrack and extended subrack.

Item	Data	Remarks

Item	Data	Remarks
Extending port of the master subrack	ETHB extending port: 0/3/0	Adding a GE remote extended subrack to a master subrack is like adding a service board to the master subrack. You can add a GE remote extended subrack in offline pre-deployment or automatic discovery mode. In offline pre-deployment mode (as in this example), an extending port on the master subrack needs to be planned for adding the extended subrack. In automatic discovery mode, an extending port does not need to be planned, and the port where the extended subrack is automatically discovered is used as the extending port.
Extended subrack ID	In offline pre-deployment mode, the extended subrack ID is 1.	The master subrack ID is always 0. After the GE remote extended subrack is introduced, the extended subrack ID starts from 1. Currently, a maximum of 32 extended subracks are supported. Therefore, the extended ID ranges from 1 to 32. In offline pre-deployment mode (as in this example), an available subrack ID needs to be specified for the extended subrack to be added. In automatic discovery mode, the master subrack automatically allocates an available subrack ID to the extended subrack that is confirmed.

Procedure

Set the auto-negotiation function of the extending port to **enable**.

```
huawei(config-if-eth-0/3)#auto-neg all enable
```

Step 1 Set the working mode of the ETHB extending board to **extend**.

```
huawei(config-if-eth-0/3)#network-role extend
huawei(config-if-eth-0/3)#quit
```

NOTICE

An ETHB board resets if its working mode switches from **extend** to other modes or from other modes to **extend**.

Step 2 Add a remote extended subrack in offline mode.

```
huawei(config)#frame add 1 MA5623AR extend-port 0/3/0
```

----End

Follow-up Procedure

After being powered on, the extended subrack automatically communicates with the master subrack to complete registration, and enters the normal state. Then, the master subrack can manage the extended subrack.

NOTE

In automatic discovery mode, perform the following operations to add a remote extended subrack:

- Run the **display network-role** command to check whether the working mode of the ETHB board is **extend**
- If the working mode is not **extend**, run the **network-role** command to set the working mode to **extend**. Note that mode switching will reset the ETHB board.
- Run the **frame confirm** command to confirm the extended subrack. Then, the extended subrack enters the normal state.

23 Voice Feature

About This Chapter

Voice communication is a method for long distance voice data transmission over networks using various technologies and protocols. Voice communication supports basic voice services, such as fax and modem services, and can be applied in residential as well as enterprise private line services.

This document covers the following contents:

- Voice feature-compliant protocols
- Voice access modes
- Voice service assurance methods

This topic describes the features in relation to the POTS interface, including basic features such as ringing and Z interface and enhanced features.

23.9 ISDN Access

The integrated services digital network (ISDN) is a CCITT standard, providing integrated transmission service for voice, video, and data. The ISDN enables the voice, video, and data to be transmitted on the data channel simultaneously.

23.10 R2 Access

R2 access enables the MA5600T/MA5603T/MA5608T to be interconnected with a private branch exchange (PBX) through the R2 signaling and helps to provide access services for users over the common twisted pairs. As a type of channel associated signaling (CAS), R2 signaling is the international standard signaling based on E1 digital networks.

23.11 FoIP

Fax over Internet Protocol (FoIP) is a fax service provided on an IP network or between an IP network and a traditional PSTN. Fax service is a data service that is widely applied on the PSTN network.

23.12 MoIP

Modem over Internet Protocol (MoIP) is a technology for providing modem services over an IP network or between an IP network and a PSTN network.

23.13 IP Z Interface Extension

IP Z interface extension is that the analog interface between an accsee device and a PBX extends to the remote place through the IP network.

23.14 Key Techniques for Improving Voice Service Quality

Voice service quality is the biggest challenge faced by the IP telephony technology. IP telephony service has a higher requirements on real-time transmission of IP packets. If IP packets are lost, or transmission delay or jitter is introduced due to transmission errors or network congestion, subscribers hear noises during calls, and even more, ongoing calls may be interrupted. The VoIP technology provides a series of technologies, such as codec, echo cancellation (EC), and voice activity detector (VAD) to improve the voice quality.

23.15 Voice Service Maintenance and Diagnosis

The maintenance and diagnosis features of voice services include these features such as the loop-line test, circuit test, call emulation test, continuity test, VBD fault diagnosis, Real-time Transport Control Protocol (RTCP) statistics and so on.

23.16 Voice Reliability

This topic describes features related to voice reliability, including dual-homing networking, highly reliable transmission (SCTP), and voice QoS.

23.17 Configuring the VoIP PSTN Service (SIP-based)

The SIP-based VoIP technology makes the transport network evolve to the IP network without decreasing the voice quality, provides more value-added functions for users, and saves expense.

23.18 Configuring the VoIP ISDN BRA Service (SIP-based)

After configuring the SIP interface, you can add VoIP ISDN BRA users on this interface to implement the VoIP ISDN BRA service.

23.19 Configuring the VoIP ISDN PRA Service (SIP-based)

After configuring the SIP interface, you can add VoIP ISDN users on this interface to implement the VoIP ISDN service.

23.20 Configuring the VoIP PSTN Service (H.248-based or MGCP-based)

This topic describes how to configure the VoIP PSTN service when the protocol adopted by the Access node is H.248 or MGCP.

23.21 Configuring the VoIP ISDN BRA Service (H.248-based)

This topic describes how to configure the VoIP ISDN BRA service on an IP network. When the Access node uses the H.248 protocol, the device supports the access of the ISDN BRA user. ISDN technology provides end-to-end (E2E) digital connection and supports multiple types of voice and non-voice telecom services.

23.22 Configuring the VoIP ISDN PRA Service (H.248-based)

This topic describes how to configure the VoIP ISDN PRA service on an IP network. When the Access node uses the H.248 protocol, the device supports the access of the ISDN PRA user. ISDN provides end-to-end (E2E) digital connection and supports multiple types of voice and non-voice telecom services.

23.23 Configuring the R2 Service

With the R2 access technology, the Access node provides access services on common twisted pair cables when interconnecting with the PBX using R2 signaling.

23.24 Configuring the H.248/MGCP-based FoIP Service

This topic describes how to configure the H.248/MGCP-based FoIP service.

23.25 Configuring the SIP-based FoIP Service

This topic describes how to configure the SIP-based FoIP service.

23.26 Configuring the MoIP Service

This topic describes how to configure the H.248/MGCP/SIP-based MoIP service for transmitting the traditional narrowband modem data service over the IP network.

23.27 Adding a POTS IP SPC

A semi-permanent connection (SPC) exclusively occupies a voice channel to meet the communication requirements and to ensure the communication quality for particular and vital access subscribers. To configure an IP SPC, configure the data such as the local IP address, local UDP port ID, remote IP address, and remote UDP port ID, and set up a direct IP connection between the two ends of the VoIP service. In this manner, the voice media data can be directly transmitted to the peer end.

23.28 Configuring the IP Z Interface Extension Service

The following network typically applies to the scenario where Z interface extension private line service needs to be carried over the IP network for the headquarters (HQ) and branch offices of an enterprise after the PSTN network reconstruction. In the following configuration example, the FXO and FXS boards are added for the Z interface extension local MSAN and remote MSAN respectively, board attributes are configured, and IP semi-permanent connections (SPCs) of the IP Z interface extension type are created between the two boards, so that users connected to the FXS board are connected to the corresponding ports on the FXO board through the SPCs.

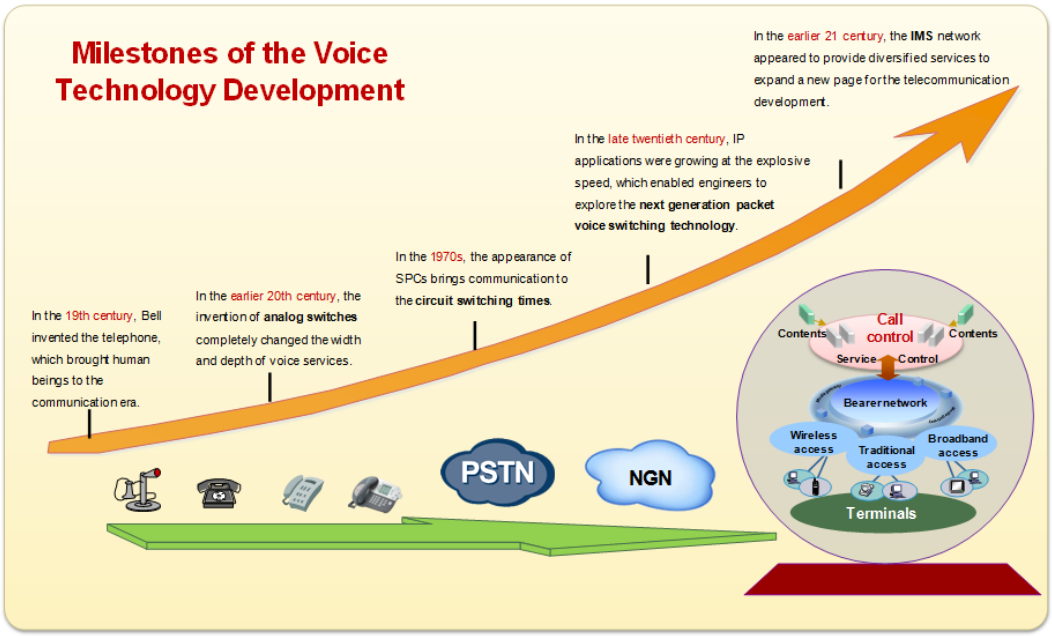
23.29 Configuring the Security and Reliability of the Voice Service

The security configuration of the voice service includes the H.248-based, MGCP-based, or SIP-based device authentication configuration, and the reliability configuration of the voice service includes the dual-homing configuration and the emergency standalone configuration.

23.1 Voice Technology Development

Figure 23-1 shows the voice technology development. The public switched telephone network (PSTN), next generation network (NGN), and IP multimedia subsystem (IMS) are widely used for communication.

Figure 23-1 Voice technology development



PSTN

PSTN is a connection-oriented network based on timeslot switching. Each circuit connection occupies a timeslot of the pulse code modulation (PCM) basic group, that is, the switching circuit is performed at the rate of 64 kbit/s, which cannot be changed. Advantages and disadvantages of PSTN are as follows:

- Advantages: Fast switching speed, accurate obtaining of call duration, short transmission delay, small jitter, supporting services with high requirements on real-time (especially telephony services)
- Disadvantages: Supporting only 64 kbit/s, exclusively occupying allocated network resources, and low resource usage

NGN

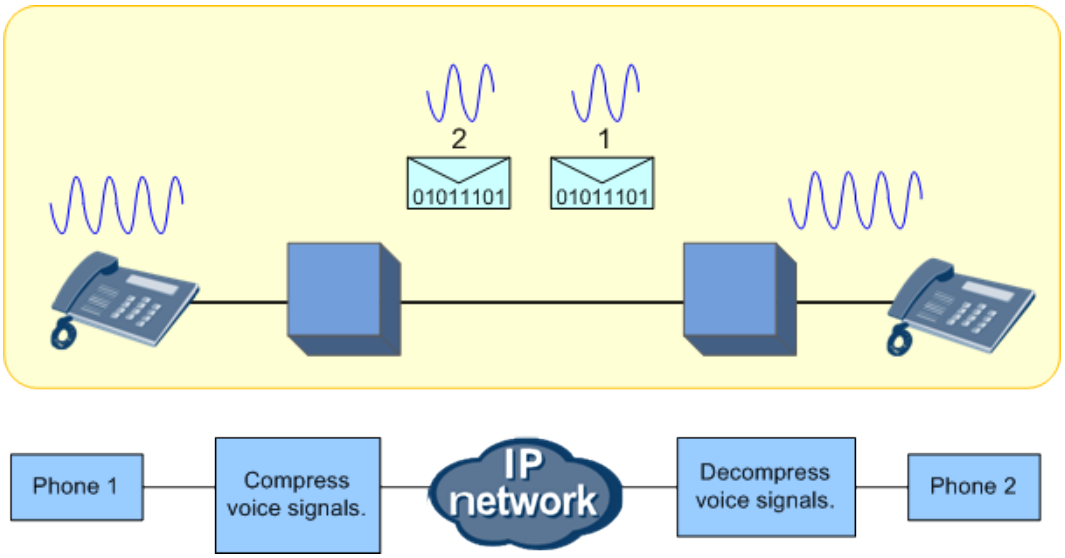
NGN is an integrated network. NGN, a packet-based network, employs the IP technology to build the carrier network and implement the separation of call control from bearing.

For packet voice services, the NGN uses a softswitch as the control layer, IP as the bearer layer, and AG as the access layer. Voice over IP (VoIP) is an important application of packet voice services.

VoIP

To implement VoIP, analog voice signals are compressed and encapsulated into data signals and then transmitted on the IP network, as shown in Figure 23-2. The example usage of VoIP is IP call. In a narrow sense, the VoIP only refers to the voice signal transmission. In a broad sense, the VoIP also refers to data signal transmission, that is, modem over IP (MoIP) and fax over IP (FoIP).

Figure 23-2 VoIP implementation



IMS Network

With the development of telecommunication technology, users want to use voice services to facilitate work and life instead of the traditional voice communication. As a result, a stronger communication platform is required to provide integrated voice, video, and mobility features, as shown in Figure 23-3. The IMS network is developed to meet requirements.

Figure 23-3 Communication requirements

The IMS is a system that controls IP-based multimedia sessions on the NGN. The IMS network contains all core NEs that implement multimedia services, including video, audio, text, and instant messaging (IM).

IP =	IP-based transmission IP-based session control IP-based service implementation
Multimedia =	Supporting multimedia services including video, audio, image, and text
Subsystem =	A system using the advanced network technology and devices

The IMS network was designed by 3rd Generation Partnership Project (3GPP) in the R5 version to support IP-based multimedia services.

The IMS network features the following:

- Same as the softswitch, call control is separated from bearing.
- Services are separated from call control, which speeds up new service provisioning, as shown in Figure 23-4.

Figure 23-4 Control, bearing, and service separation

- The IP-based IMS network is an integrated core network that can be shared by the mobile and fixed networks.
- The IMS network uses E2E SIP signaling. Services and terminals are developed toward intelligence.
  - The control plane using the SIP protocol to control signaling in a centralized manner.
  - The service plane using the SIP protocol to provide a uniformed session mechanism for all services.

23.2 Voice Service Networking Applications

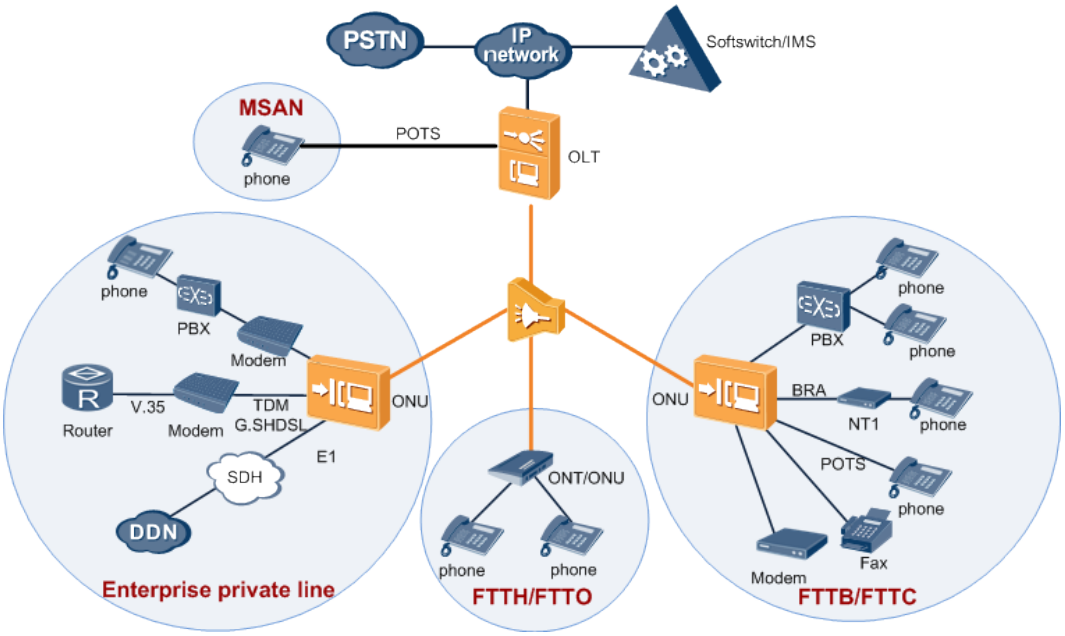
Voice services, including POTS, fax, modem, ISDN, and R2 services, apply to multiservice access node (MSAN), fiber to the building (FTTB), fiber to the curb (FTTC), fiber to the home (FTTH), fiber to the office (FTTO), and enterprise private line scenarios.

Networking Applications

An access gateway (AG) supports the following functions:

- Complies with SIP, H.248, or MGCP, and works with the softswitch or IMS to support the VoIP service.
- Supports ISDN BRA and PRA services.
- Supports the R2 service over E1 lines.
- Supports fax (FoIP) and modem (MoIP) services.
- Supports the TDM SHDSL service using V.35 or E1 upstream transmission, reconstructing traditional voice networks. Compared with the V.35 and E1 services, the TDM SHDSL service supports a longer transmission distance.
  - Prolonged E1 transmission distances: The TDM SHDSL modem on the user side connects to the PBX using an E1 (ISDN PRI) interface, and the modem connects to the AG in TDM G.SHDSL access mode. Then, the AG sends signaling streams to the IP network and exchanges voice service flows with other voice devices using a media gateway (MG).
  - Prolonged V.35 transmission distances: The TDM SHDSL modem on the user side connects to the user-side device using a V.35 (N x 64 kbit/s private line) interface, and the modem connects to the AG in TDM SHDSL access mode. Then, the AG sends data to a DDN network using an SDH network, implementing N x 64 kbit/s DDN private line access.

Figure 23-5 Voice service networking

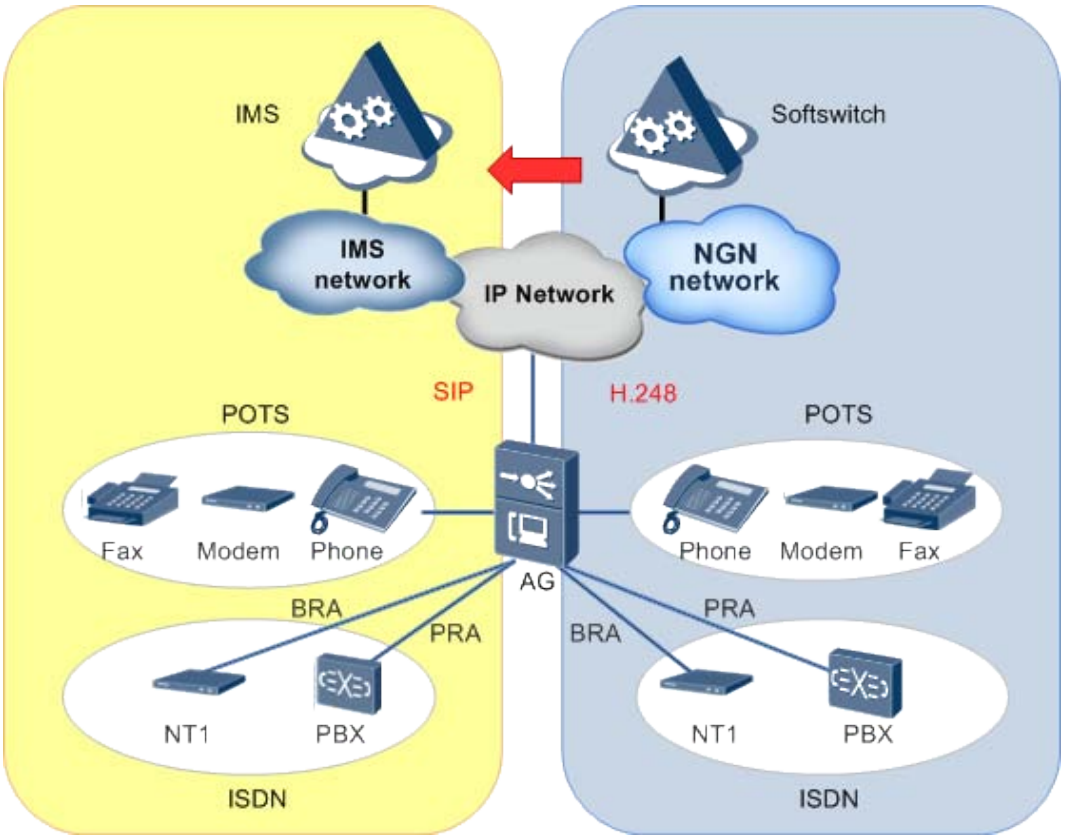


SIP and H.248 Dual-Upstream Transmission

An AG supports both SIP and H.248 upstream transmission. This upstream mode supports smooth migration from a softswitch network to an IP multimedia subsystem (IMS) network, as shown in Figure 23-6. The SIP and H.248 dual-upstream transmission applies to network reconstructions.

- IMS networks provide more value-added services. For the users to be able to migrate from softswitch to IMS, the AG can switch the upstream transmission mode from H.248 to SIP, ensuring that the migration does not interrupt other users' services.
- Each virtual access gateway (VAG) can be configured to support H.248 or SIP.
- The maximum number of users supported by an AG with both H.248 and SIP enabled is the same as that supported by the AG with either of the protocols enabled.

Figure 23-6 Dual-upstream transmission networking



23.3 Voice Feature Overview

Access devices support the following basic voice features to help carriers provide high-quality voice services.

Table 23-1 Basic voice features

Basic Feature	Description
Protocol Voice protocols are used for communication between an access device and an upper-layer gateway control device. The access device supports H.248, Media Gateway Control Protocol (MGCP), and Session Initiation Protocol (SIP). You can run the **display protocol support** command to query the protocol used on the device. If the protocol used on the device is different from the planned protocol, run the **protocol support** command to switch the protocol.	23.5 SIP Voice Feature 23.7 H.248 Voice Feature 23.6 MGCP Voice Feature
Access Mode	23.9 ISDN Access

Basic Feature	Description
A voice access mode is based on terminal type and service requirements. The access device supports POTS, ISDN, and R2 access modes. Voice communication also supports fax and modem services. • A POTS network is a connection-oriented circuit switched network based on timeslot switching. Each circuit connection uses a timeslot in a pulse code modulation (PCM) group. That is, the circuit switched rate is 64 kbit/s. • The ISDN access can be basic rate access (BRA) or primary rate access (PRA). The BRA access provides 2 B channels and 1 D channel. The rates of B and D channels are 64 kbit/s and 16 kbit/s, respectively. The PRA access provides 30 B channels and 1 D channel. The rates of B and D channels are both 64 kbit/s. • In R2 access mode, the access device connects to a private branch exchange (PBX), which communicates with the access device through R2 signaling.	23.8 POTS Access 23.10 R2 Access 23.11 FoIP 23.12 MoIP
Key Techniques for Improving Voice Service Quality The VoIP technology provides a series of technologies, such as codec, echo cancellation (EC), and voice activity detector (VAD) to improve the voice quality.	23.14 Key Techniques for Improving Voice Service Quality

Table 23-2 Voice service features

Service Type	Description
VoIP The VoIP service uses the IP packet switched network for transmission after the traditional analog voice signals are compressed and packetized, to lower the cost of the voice service.	For the SIP protocol: • SIP-based VoIP • 23.17 Configuring the VoIP PSTN Service (SIP-based) • 23.18 Configuring the VoIP ISDN BRA Service (SIP-based) • 23.19 Configuring the VoIP ISDN PRA Service (SIP-based) • 23.23 Configuring the R2 Service
	For the H.248 protocol: • H.248-Based VoIP • 23.20 Configuring the VoIP PSTN

Service Type	Description
	Service (H.248-based or MGCP-based) • 23.21 Configuring the VoIP ISDN BRA Service (H.248-based) • 23.22 Configuring the VoIP ISDN PRA Service (H.248-based) • 23.23 Configuring the R2 Service
	For the MGCP protocol: • MGCP-Based VoIP • 23.20 Configuring the VoIP PSTN Service (H.248-based or MGCP-based)
MoIP The modem over Internet Protocol (MoIP) refers to the modem service provided on the IP network or between the IP network and the traditional PSTN network.	For the SIP protocol: • SIP-Based MoIP • 23.26 Configuring the MoIP Service
	For the H.248 protocol: • H.248-Based MoIP • 23.26 Configuring the MoIP Service
	For the MGCP protocol: • MGCP-Based MoIP • 23.26 Configuring the MoIP Service
FoIP The fax over Internet Protocol (FoIP) refers to the fax service provided on the IP network or between the IP network and the traditional PSTN network.	For the SIP protocol: • H.248-Based FoIP • 23.25 Configuring the SIP-based FoIP Service
	For the H.248 protocol: • H.248-Based FoIP • 23.24 Configuring the H.248/MGCP-based FoIP Service
	For the MGCP protocol: • MGCP-Based FoIP • 23.24 Configuring the H.248/MGCP-based FoIP Service
Line hunting Line hunting is a feature that allows a group of ports to share a group of called party numbers by specifying a hunting group and hunting policy. Only the SIP protocol supports this feature.	• Line Hunting • 23.17.3 (Optional) Configuring Line Hunting

Service Type	Description
POTS IP SPC To configure an IP SPC, configure the data (including the local IP address, local UDP port, remote IP address, and remote UDP port), set up an IP direct connection between the two ends of the voice service. In this manner, the voice media data can be directly transmitted to the peer end.	• 23.8.5 POTS IP SPC • 23.27 Adding a POTS IP SPC

Table 23-3 Voice security features

Security	Description
Device authentication Device authentication is a method to improve the security of the core network and prevent illegal devices from registering with the core network device.	For the SIP protocol: • Configuring Device Authentication Based on SIP
	For the H.248 protocol: • Configuring Device Authentication (H.248-based)
	For the MGCP protocol: • Configuring Device Authentication (MGCP-based)
Dual-homing Dual homing is an NGN (Next Generation Network) total solution. Based on this solution, when the active softswitch or the link from the MG to the active softswitch is faulty, the MG need be switched to the standby softswitch immediately to prevent call services of users connected to the softswitch and the MG from being affected.	For the SIP protocol: • 23.16.4 SIP Dual Homing • Configuring the SIP-based Dual Homing
	For the H.248 protocol: • 23.16.1 H.248/MGCP Dual Homing • Configuring H.248-based Dual Homing (Multi-homing)
	For the MGCP protocol: • 23.16.1 H.248/MGCP Dual Homing • Configuring MGCP-based Dual Homing
Multi-homing As an enhancement of dual-homing, multi-homing is a configuration in which a media gateway (MG) is homed to the primary media gateway controller (MGC), secondary MGC, and disaster-recovery MGC.	• 23.16.2 H.248 Multi-homing • Configuring H.248-based Dual Homing (Multi-homing)
Emergency standalone Emergency standalone is a solution in which the	• 23.16.3 Emergency Standalone • 23.29.2 Configuring Inner

Security	Description
users on the same MG can call each other even when the interface between the MG and the softswitch is interrupted.	Standalone (H.248-based or SIP-based)

Table 23-4 POTS port maintenance and test features

Troubleshooting Method	Description
A POTS user loop line test is used to test the electrical indicators of the line from the test device (an access node) to a phone. When users' POTS services are faulty, POTS user loop line tests can be performed to test the performance and electrical indicators of the loop line to diagnose whether the loop line is faulty.	23.15.2 POTS User Loop Line Test
A POTS user circuit test is used to check whether the chip of a POTS board functions normally. If the POTS services are faulty and the loop line works normally, POTS user circuit tests can be used to test the functions (such as the ringing and power feeding) and some parameters (such as the feeding voltage and ringing voltage) of the board circuit to check whether the circuit works normally.	23.15.3 POTS User Circuit Test
A POTS port loop test is used to test the hardware and configurations related to POTS services during device installation or before POTS service provisioning. It helps reduce the number of site visits and minimize maintenance costs.	23.15.4 POTS Port Loop Test
A search tone test is a simple line fault locating function intended for maintenance engineers. In a search tone test, the test module sends voice signals with the specific frequency and amplitude to a line, and then maintenance engineers use a receiver or a dedicated device to detect the signals on the line. In addition, search tone tests can help maintenance engineers pinpoint the specific line among multiple user lines.	23.15.5 Search Tone Test
In a signal tone test, the system sends the signal tone signals to a specific port of a POTS board and makes the port loop back the signals, and then checks whether the loopback signals can be detected. This test function helps maintenance engineers check whether the system can normally process the detection of the user off-hook and signal tone and locate hardware	23.15.6 Signal Tone Test

Troubleshooting Method	Description
faults related to the user off-hook and signal tone playing.	
A call emulation test emulates call functions to verify data configuration for the voice service. The call emulation test can also be used to locate voice service faults.	23.15.1 Call Emulation Test

23.4 Basic Concepts in Voice Services

Learning these basic concepts facilitates deep understanding of voice services.

23.4.1 Voice Media and Signaling

Media and signaling play important roles in voice services.

- Voice signaling: Used to set up and control voice communication between two telecommunication entities. Different from IP protocols, signaling protocol fields carry commands. Common signaling protocols are MGCP, H.248, and SIP.
- Voice media: Used to carry and normally transmit voice communication contents.

Voice Media

RTP

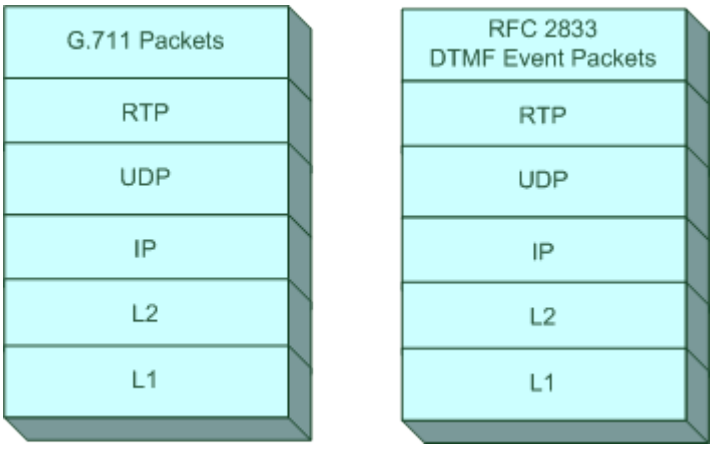
Real-Time Transport Protocol (RTP) is dedicated for multi-media streams over the Internet. In a VoIP network, RTP carries media streams. For details about RTP, see the *RFC3550*.

RTP provides end-to-end (E2E) services to transport real time data (including audios and videos) in the format defined in G.711, RFC2833, and RFC2198. Figure 23-7 shows the format of an RTP packet.

Figure 23-7 RTP packet format

RTP runs on top of User Datagram Protocol (UDP) to make use of its multiplexing and checksum services. However, RTP may be used with other suitable underlying network or transport protocols. Figure 23-8 shows the RTP protocol stack model.

Figure 23-8 RTP protocol stack model



RTP receives media streams from the upper-layer device and encapsulates the streams into RTP packets. Then RTP sends the packets to the lower-layer device. The lower-layer protocol transmits RTP and Real-Time Transport Control Protocol (RTCP) packets through different ports. For example, if UDP is used as the lower layer protocol, the protocol uses a port with the ID of an even number to transmit RTP packets and uses the port with the ID of the odd number following the even number to transmit RTCP packets.

RTCP

RTP itself ensures real-time data transmission, but cannot provide a mechanism for reliably transmitting data in sequence or a traffic and congestion control mechanism. It provides the mechanisms using RTCP.

RTCP is based on the periodic transmission of control packets to all participants in the session, using the same distribution mechanism as the data packets, so that it provides feedback on the quality of the data distribution.

- RTCP packets take 5% of bandwidths.
- RTCP packets contains ring delay, packet loss stream, and jitter for QoS monitoring.

Figure 23-9 shows the format of an RTCP packet.

Figure 23-9 RTCP packet format

RFC 2833

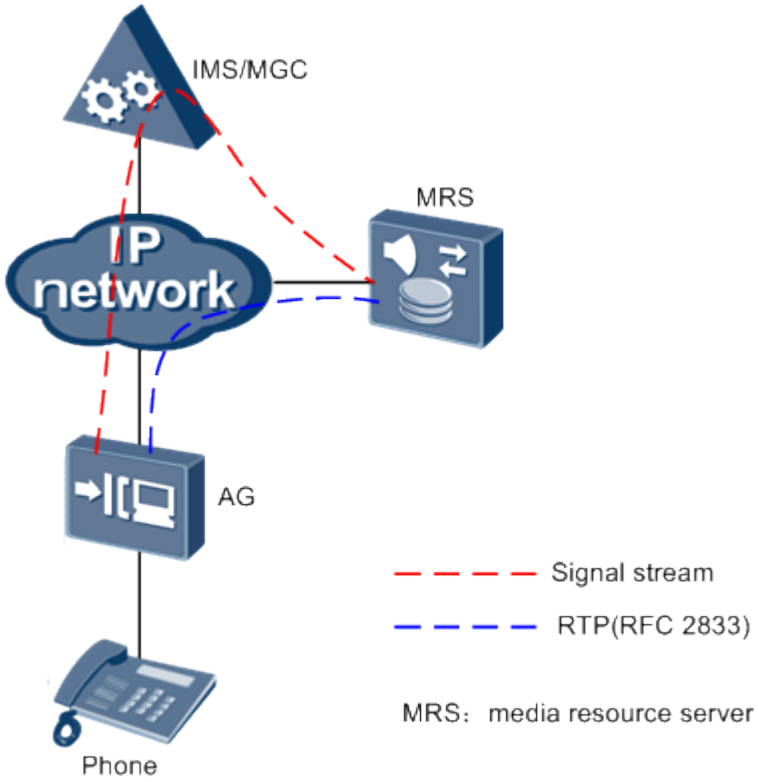
RFC2833 defines a dedicated format to reliably transmit important data, such as signal tone, event, dual tone multi-frequency (DTMF) signal in G.711, G.723, and G.729 communication.

Figure 23-10 shows the typical RFC2833 application scenario.

1. After the user dial a number, the softswitch controls the access gateway (AG) to create and transmit RTP voice media streams to the media resource server (MRS).

2. The MRS plays an announcement of dialing the number to the AG.
3. The user dials the number. Then the number is transmitted to the MRS through the RFC2833 packet that is carried over voice media RTP.
4. The MRS collects and transmits the number to the softswitch.

Figure 23-10 Typical RFC2833 application scenario



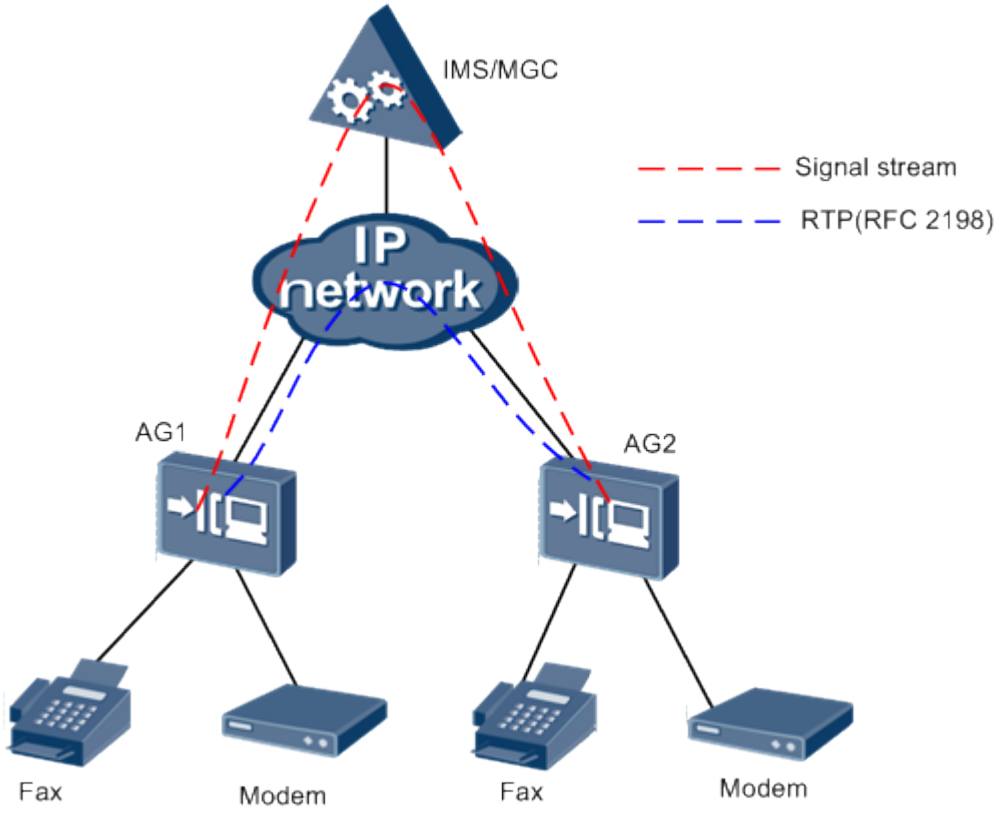
RFC2198 Redundancy

RFC2198 describes the RTP payload format for redundant audio data, which can be used for the RFC2833 digit collecting, fax transparent transmission service, and modem transparent transmission service.

NOTE
RFC2198 redundancy is unnecessary for T.38 fax services, this is because T.38 fax service supports redundancy for its own.

RFC2198 improves the reliability of data transmission through redundant transmission. When the network quality is poor, redundant transmission can ensure the service quality and reduce impacts brought by distorted signals. Figure 23-11 shows RFC2198 redundancy application

Figure 23-11 RFC2198 redundancy application



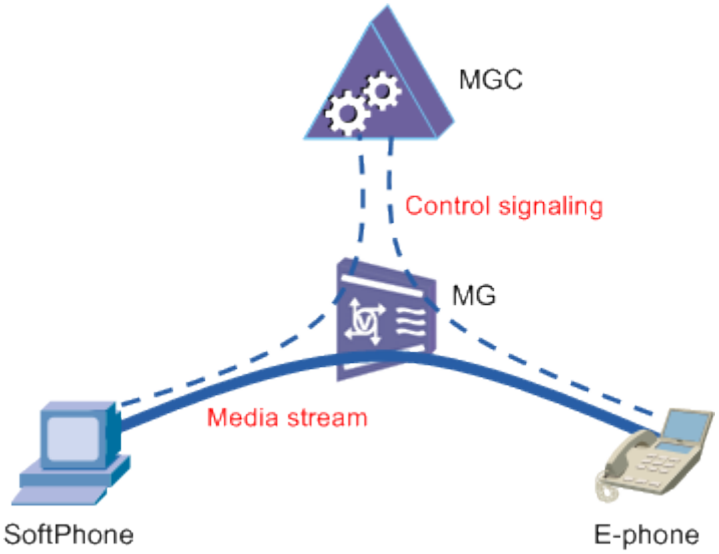
Voice Signaling

The signaling technology implements phone calls. The commonly used VoIP control signaling systems contain MGCP, H.248, and SIP.

MGCP Protocol

Media Gateway Control Protocol (MGCP) is defined in the RFC2705 standard and it defines a call control structure in which call control is separated from service bearer. As shown in Figure 23-12, call control is separate from the media gateway (MG) and is processed by the media gateway controller (MGC). Therefore, MGCP is in nature a master-slave protocol. The MG establishes various service connections under the control of the MGC.

Figure 23-12 MGCP Master-Slave Control



H.248 Protocol

H.248 is the same type of protocol as MeGaCo and completed by the ITU-T and IETF together, used as a media gateway control protocol between an MGC and an MG. It takes the place of MGCP. H.248 features the following:

- Functions on the basis of MGCP and therefore it inherits all advantages of MGCP.
- Works in master-slave mode.
- Uses binary coding or text coding for H.248 messages. MGC must support these two coding modes and MG supports either of them.
- Uses User Datagram Protocol (UDP), Transmission Control Protocol (TCP) or Stream Control Transmission Protocol (SCTP) (IP-based signaling transmission) for underlying transmission.

Compared with MGCP, H.248 has the following advantages:

- Supports voice services and multi-media connections.
- Supports text coding and binary coding.
- Features expandability.

SIP Protocol

Session initiation (SIP) is a session control protocol running at the application layer, which sets up, modifies, and terminates a session. A session can be an application, such as multi-media conference and Internet call.

Comparison Between the H.248 Protocol and SIP Protocol

Table 23-5 describes comparison between the H.248 protocol and SIP protocol.

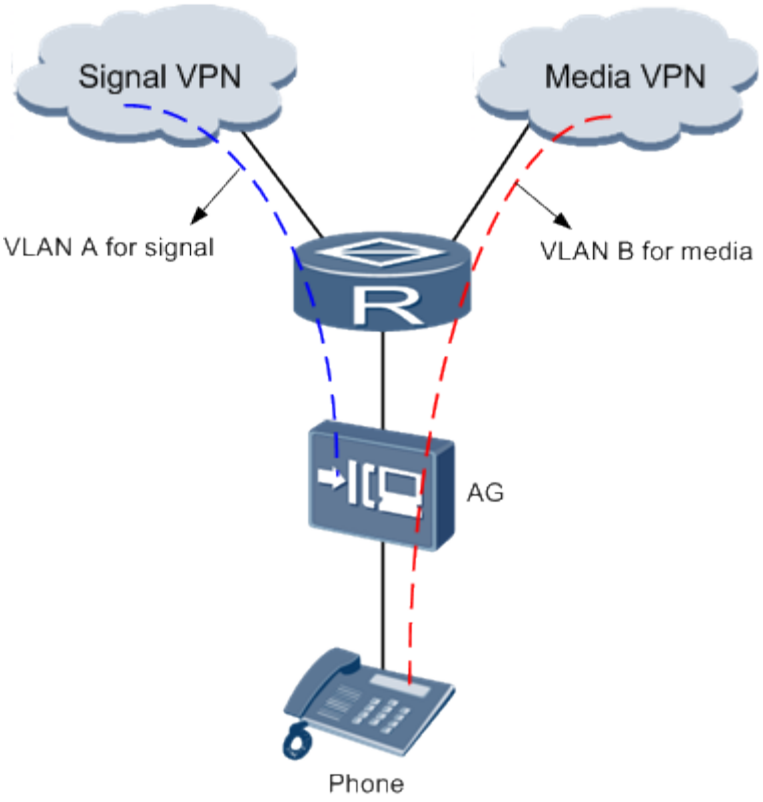
Table 23-5 Comparison Between the H.248 protocol and SIP protocol

Item	SIP Protocol	H.248 Protocol
Standard	IETF/TISPAN	ITU_T/TISPAN
Architecture	Distributed/Intelligent clients	Centralized/Dumb end_point
Call control	Proxy/Redirect Server	Call agent/MGC
Transport protocol	UDP/TCP/SCTP	UDP/TCP/SCTP
Multi-media service supported	Yes	Yes
Supplementary service	Provided by endpoints or by call control	Provided by call control
ISDN service	Not defined in TISPAN R1	Use IUA Support

Separation of Media and Signaling Streams

Separation of media and signaling streams indicates that signaling streams (H.248/SIP) and media streams can be transmitted upstream to different virtual private networks (VPNs) through different IP addresses and VLANs. The separation facilitates network planning and meets control requirements, as shown in Figure 23-13.

Figure 23-13 Separation of media and signaling streams



Application scenarios of separation of media and signaling streams:

- Signaling and media flows are transmitted upstream to different VPNs and they use different control policies, such as QoS.
- Different user groups use different signaling IP addresses/VLANs and media IP addresses/VLANs. This meets the special monitoring requirements.

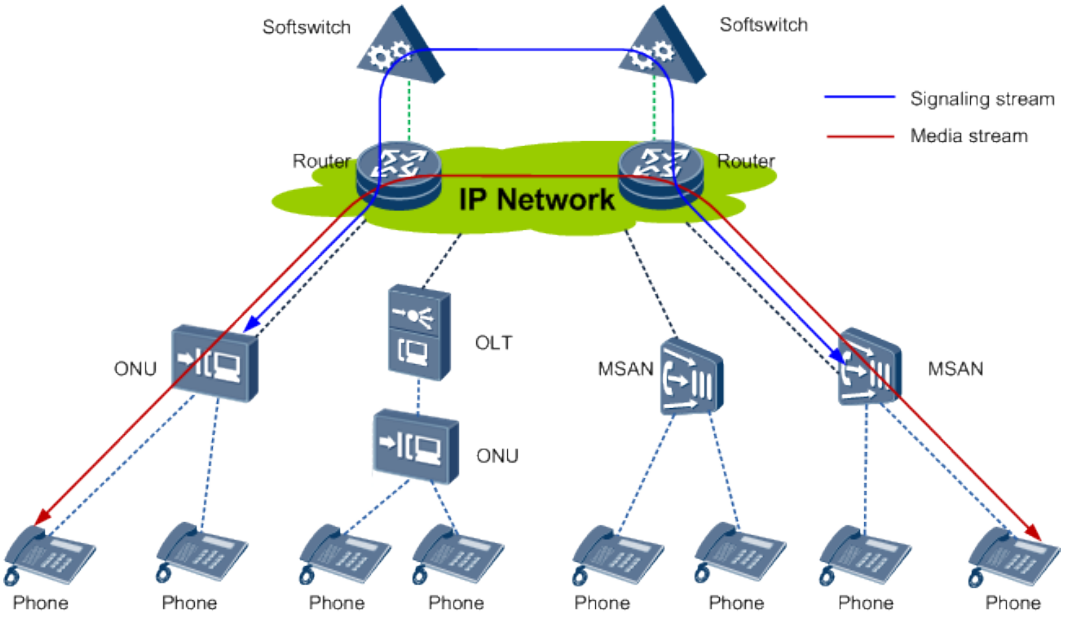
External Direction of Voice Media Streams and Signaling Streams

Voice services can be classified into the following 3 scenarios in which media streams and signaling streams are transmitted in different directions.

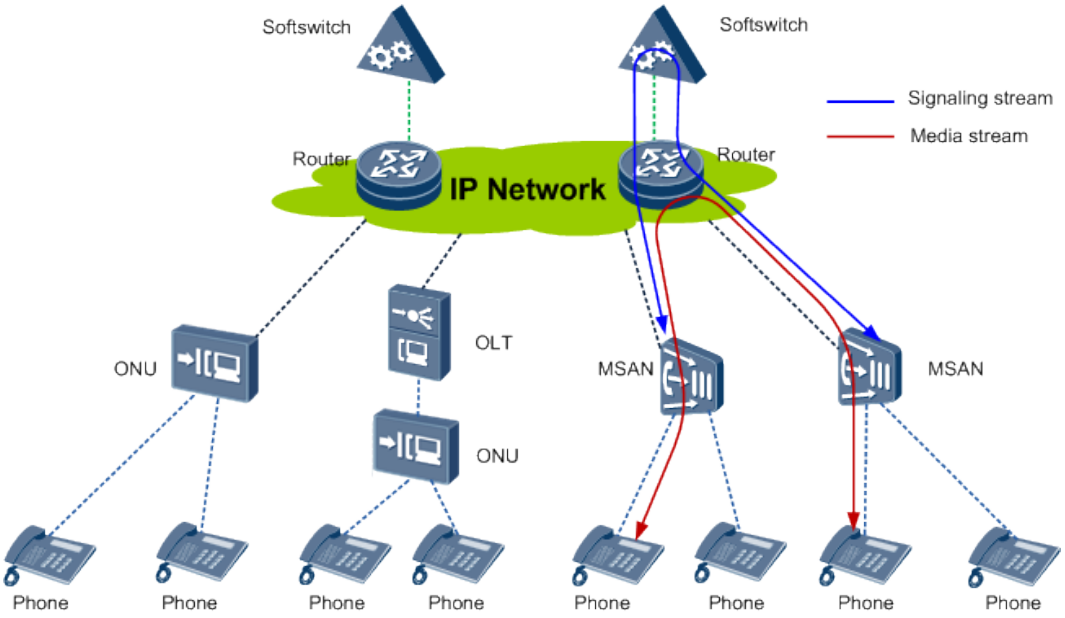
- Communication under different softswitches
- Communication under the same softswitch but different access gateways (AGs)
- Communication under the same softswitch and same AG

The following shows directions of voice media streams and signaling streams in these 3 scenarios:

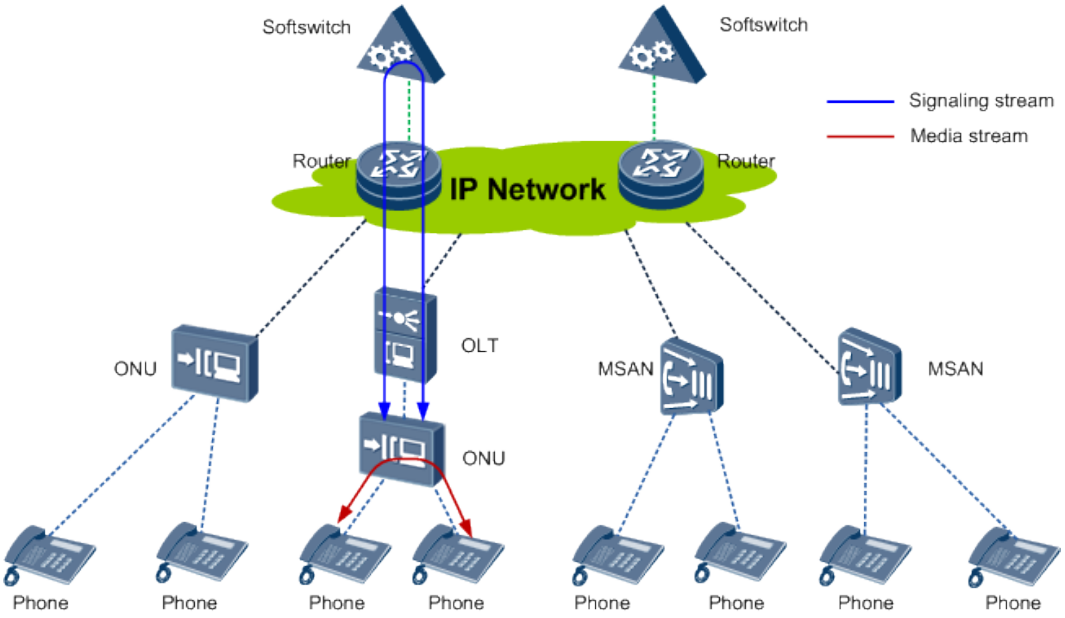
Communication Under Different Softswitches



Communication Under the Same Softswitch but Different AGs



Communication Under the Same Softswitch and Same AG



Internal Direction of Voice Media Streams and Signaling Streams

POTS Access Mode

The following flash video demonstrates the directions of media streams and signaling streams inside an AG device in POTS access mode.

Functions of each module are as follows:

- The subscriber line interface circuit (SLIC) processes analog signals, including feeding phone sets, sending voice frequencies to phone sets, and generating ringing, as well as detecting off-hook, pulse dialing, on-hook, and hookflash signals.
- The coder/decoder (CODEC) converts between analog and digital signals. It converts analog signals to digital signals in the upstream direction and digital signals to analog signals in the downstream direction.
- The digital signal processor (DSP) supports the following functions:
  - Codes and decodes voice signals. The DSP encapsulates the digital signals sent by the CODEC into VoIP packets in the upstream direction and restores the VoIP media streams transmitted over the GE bus to digital signals in the downstream direction.
  - Manages the SLIC and CODEC using the SPI bus.
- The GE LSW and voice processing module process signaling streams and media streams. They determine whether to discard or forward signaling/media streams based on the IP address of the signaling/media streams and the UDP port number. The voice processing module is a daughter board on the control board.

ISDN Access Mode

The following flash video demonstrates the directions of media streams and signaling streams inside an AG device in ISDN access mode.

Functions of each module are as follows:

- Network termination (NT) 1: functions similarly as the physical layer of the OSI reference model. The functions of NT1 are associated with inherent physical and electrical characteristics of the network.
- Terminal adapter (TA): Non-ISDN terminals do not support channel-D functions. To connect a non-ISDN terminal to the ISDN network, a TA is required.
- ISDN chip group: connects to NT1 using a U interface for ISDN BRA services. The ISDN chip group converts analog signals to digital signals and codes and decodes a U interface (such as 2B1Q and 4B3T codes).
- E1 chip group: connects the E1 interface to NT 1 for the ISDN PRA service. It converts analog signals to digital signals.
- DSP: codes and decodes voice signals. The DSP encapsulates the digital signals sent by the ISDN or E1 chip group into VoIP packets in the upstream direction and restores the VoIP media streams transmitted over the GE bus to digital signals in the downstream direction.
- The GE LSW and voice processing module forward signaling streams and media streams. They determine whether to discard or forward signaling/media streams based on the IP address of the signaling/media streams and the UDP port number. The voice processing module is a daughter board on the control board.

23.4.2 VAG

The virtual access gateway (VAG) is a solution in which one AG device can be simulated into multiple AG devices. Different VAGs can connect to different IMSs/softswitches, as shown in Figure 23-14. The purposes of the VAG are as follows:

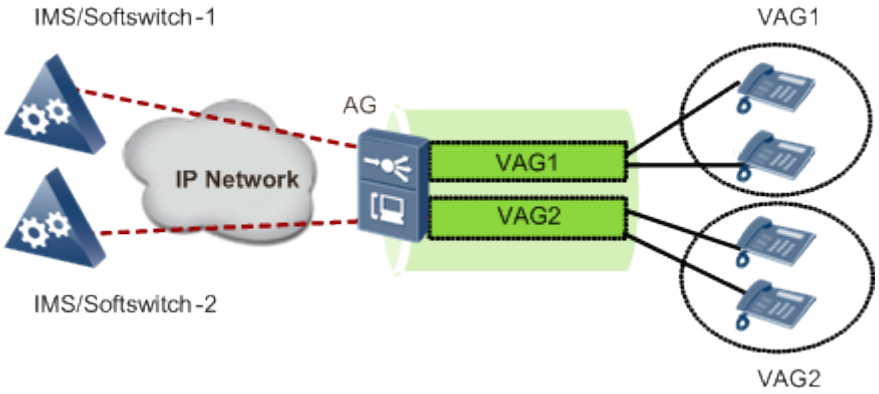
- Differentiated service

  Simulate multiple logical AGs on a physical AG, and enable different logical AGs to connect to different customer groups. In this way, different convergence rates are provided for different customer groups to meet the identified service requirements.
- Virtual operation

  In the AG networking application, certain small-scale operators do not buy independent AG devices, but rent the AG devices from other operators to provide the service. For the renters, they may rent out a device to multiple operators to improve the device utilization rate. Simulate multiple logical AGs on a physical AG, and enable different logical AGs to connect to different softswitches. In this way, the wholesale service requirements can be met.
- Sharing load on the IMS/softswitch

  One IMS/softswitch has limited capacity. If subscribers connected to different VAGs are distributed to multiple IMS/softswitches, the workload on one IMS/softswitch becomes lighter.

Figure 23-14 VAG networking



The VAG call process and the traditional call process are similar. The main differences are as follows:

- The DSP distributes resources. After a subscriber picks up the phone, the AG applies for DSP resources. If DSP resources are assigned to the VAG, the VAG obtains DSP resources from the shared resource pool. If DSP resources are not assigned to the VAG, the VAG obtains DSP resources from the exclusive resource pool.
- The IP addresses in the IP packets during the connections and conversations of the subscribers under the same VAG are the media IP address corresponding to the VAG.

Each VAG can be configured and managed independently. For the softswitch, each VAG is an independent AG, and can be configured with related attributes separately, such as authentication, ringing mapping, and terminal layered mode. When configuring a subscriber, specify a VAG for the subscriber. Subscribers under different VAGs can share the same terminal ID. The total number of subscribers configured on all VAGs cannot exceed the maximum number configured in the system.

To configure a VAG, add a VAG interface (MG interface), and enter the VAG interface mode to configure interface parameters by running commands. Each VAG can be configured with independent signaling IP address and media stream IP address (the signaling IP address and media stream IP address in the same VAG can be different) in different VLANs.

23.4.3 Local Digitmap

A digitmap is a dialing scheme configured on an AG. The AG collects digits dialed by calling parties based on the digitmap. With the digitmap, the AG reports a group of digits each time, reducing the number of signaling exchanges between the AG and the softswitch, thereby improving efficiency. The digitmap configured on the AG is called local digitmap. An H.248-compliant AG can use the local digitmap or digitmap issued by the softswitch. A SIP-compliant AG must use the local digitmap

Definition

After a subscriber dials a number, the AG matches digits of the number. If the matching is successful, the AG sends the collected digits to the softswitch. The softswitch analyzes the number and corresponding services. If the matching fails, the AG discards the digits.

On the NGN, terminals send called numbers to the softswitch on the control layer through the AG on the access layer. Numbers can be transmitted in in-band or out-band mode.

- In-band transmission occupies IP media stream resources. In this mode, the AG transfers numbers in RTP or RFC 2833 format to the upper layer network. Such processing is not related to digitmaps.
- In out-band transmission mode, the AG transfers numbers through signaling streams. The out-band transmission has the following two modes:
  - Reporting a number digit by digit: The AG reports a digit to the softswitch after a subscriber dials a digit. The number of signaling exchanges between the AG and softswitch is determined by the number of digits contained in a dialed number. Such processing is not related to digitmaps.
  - Reporting a complete number: The AG reports a complete number dialed by the subscriber to the softswitch using one signaling message. This type of processing lowers the burden on the softswitch and is preferentially selected. To report a complete number, the AG must know when to collect digits and when to report digits. The digitmap determines whether dialed digits are valid.

Digitmap Example

The following digitmap is used as an example:

[2-8]xxxxxxx | 13xxxxxxxxx | 0xxxxxxxxx | 9xxxx | 1[0124-9]x | * | # | x.# | [0-9*#].T

This digitmap consists of nine character strings.

- The first character string "[2-8]xxxxxxx" indicates that the matched number must contain eight digits, the first digit must be one of 2 to 8, and remaining seven digits can be any of 0 to 9.
- The second character string "13xxxxxxxxx" indicates that the matched number must contain 11 digits, the first digit must be 1, the second digit must be 3, and remaining nine digits can be any of 0 to 9.
- The third character string "0xxxxxxxxx" indicates that the matched number must start with 0 and contain 10 digits.
- The fourth character string "9xxxx" indicates that the matched number must start with 9 and contain 5 digits.
- The fifth character string "1[0124-9]x" indicates that the first digit must be 1, the second digit cannot be 3, and the third digit can be any of 0 to 9.
- The sixth character asterisk (*) indicates that the matched character must be an asterisk (*).
- The seventh character pound sign (#) indicates that the matched character must be a pound sign (#).
- The eighth character string "x.#" indicates that the AG matches one of 0 to 9 for multiple times or even does not match any digits and stops the matching until detecting a pound sign (#).
- The ninth character string "[0-9*#].T" indicates that the AG starts a timer, matches digits 0 to 9 for multiple times or even does not match any digits, asterisk (*), or pound sign (#), and stops the matching only after the timer expires.

NOTE

The dot (.) indicates zero or multiple times of matching. However, at least one character must be matched for this character string. Therefore, at least one of 0-9, *, and # must be matched.

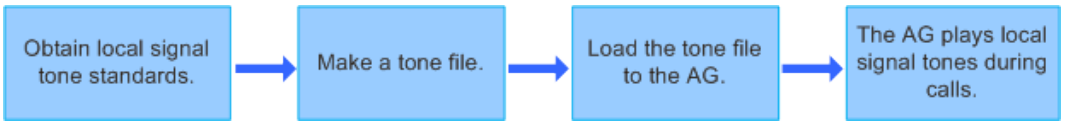
23.4.4 Local Tone

The softswitch or AG can play tones to terminals. Tones played by the AG are called local tones. Before enabling the AG to play local tones, load a tone file to the AG. For parameter tones, configure them on the AG.

Tone File

Signal tone standards vary depending on country requirements. As shown in Figure 23-15, customize a tone file, load the tone file to the AG, and the AG plays the customized tones, such as dial tone, ring back tone, and busy tone during calls.

Figure 23-15 Tone played by the AG using the tone file



Voice files, generally named **voice.efs**, are stored in the flash memory of a control board. A voice file describes the tone playing types supported by the DSP. The description covers the signal tone type, frequency, duration, and level. After the system is initialized, tone playing parameters are configured on the DSP. When the system requests to play a tone for a subscriber, the DSP reads the configuration and generates the desired signal tone on a real-time basis.

The signal tones specified in a voice file can be parameter tones, waveform tones, or announcements.

- Parameter tones are simple tones, including dialing tones, busy tones, and ring back tones. The system issues parameter tone attributes, such as the frequency, energy, duration, and cadence, to the DSP. The DSP then generates parameter tones.
- Waveform tones are simple tones. The system records these tones into a voice file and stores the file in the DSP. When the system needs to play a waveform tone for a subscriber, the DSP plays the recorded voice data for the subscriber.
- Announcements are audio messages played to subscribers, such as "The subscriber you dialed is busy. Please call later." The system records announcements into a voice file and stores the file in the DSP. When the system needs to play an announcement for a subscriber, the logic or DSP plays the recorded voice data for the subscriber.

In addition to voice file recording, the system supports parameter tone customization. A customized parameter tone takes effect in the next tone playing, which does not require service board resetting.

Customized Parameter Tones

Users can customize a parameter tone by specifying a series of cadences and tone playing rules. Each cadence defines the signal frequency, energy, and duration.

In a parameter tone, cadences can be played in one of the following ways:

- Sequential play: Each cadence is played once in sequence according to the break-make ratio.

- Continuous play: All cadences are cyclically and continuously played according to the break-make ratio.
- Cyclic play: All cadences are cyclically played and the number of cycles can be defined.

NOTE

In a parameter tone, continuous cadences can be cyclically played. For example, in Figure 23-16, special dial tone cadences 1 and 2 can be continuously played. When configuring a cadence cycle, ensure that the following requirements are met:

- The cycle end cadence must be greater than the cycle start cadence. The start and end cadences must have been configured.
- If multiple cycles are configured, the start cadence of the next cycle must be greater than or equal to the end cadence of the previous cycle. In addition, the cadences in each cycle cannot overlap or nest. For example, a parameter tone contains eight cadences, cadence 1 through cadence 8, and cycle 1 contains cadences 1, 2, and 3. Then, the start cadence of cycle 2 must be greater than or equal to cadence 3 and cannot contain cadence 1 or cadence 2.

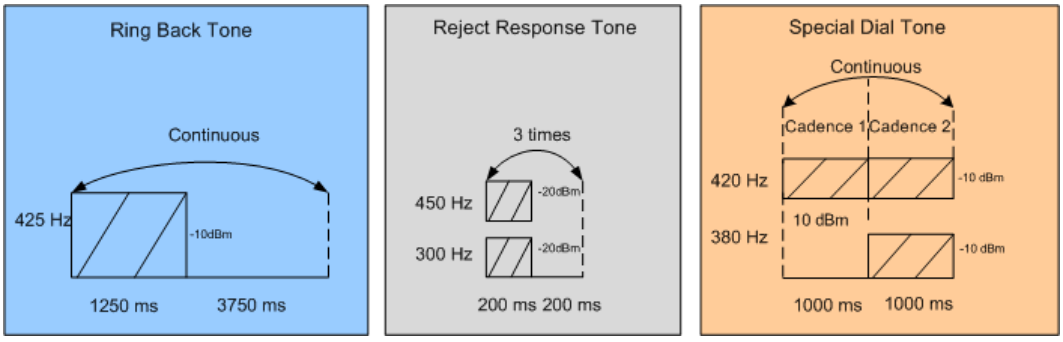
A cadence contains the following items:

- Frequency: A cadence can contain only one frequency, two frequencies, three frequencies, or four frequencies.
- Energy: The energy obtained when the POTS gain is 0 dBm and the line length is 0 km is used in this section.
- Play mode: A cadence can be played only once or repeatedly played according to the break-make ratio. If the cadence is configured to repeatedly play, the number of plays is not limited.

Figure 23-16 shows ring back tone, reject response tone, and special dial tone diagrams.

- A ring back tone is continuously played with a single cadence. The cadence contains one frequency, which is 425 Hz. The energy is -10 dBm. The cadence is played according to the break-make ratio. That is, the cadence is played for 1250 ms and stopped for 3750 ms.
- The reject response tone is cyclically played with a single cadence. The number of cycles is 3. The cadence contains two frequencies, which are 300 Hz and 450 Hz, respectively. The energy is both -20 dBm. The cadence is played according to the break-make ratio. That is, the cadence is played for 200 ms and stopped for 200 ms.
- The special dial tone is continuously played with two cadences, cadence 1 and cadence 2.
  - Cadence 1 contains one frequency, which is 420 Hz. The energy is -10 dBm. The cadence is played only once and the duration is 1000 ms.
  - Cadence 2 contains two frequencies, which are 380 Hz and 420 Hz, respectively. The energy is both -10 dBm. The cadence is played only once and the duration is 1000 ms.

Figure 23-16 Example of generating parameter tones



Feature Dependencies and Limitations

- Users only can customize parameter tones. If the AG is required to play a waveform tone or announcement, a voice file is required.
- If a parameter tone is both customized and configured using a voice file, the customized configuration preferentially takes effect.

23.4.5 Accounting

This section describes accounting on pay phones using coins or IC cards.

Introduction

The AG supports three types of accounting: polarity reversal accounting, 12/16KC accounting, and polarity reversal pulse accounting.

- Polarity reversal accounting: The pay phone starts to account immediately after detecting a voltage reversal between A and B wires.
- 12/16KC accounting: It is also called KC accounting. The pay phone performs accounting based on 12 kHz or 16 kHz high frequency signals sent by the POTS board.
- Polarity-reversal pulse accounting: The pay phone performs accounting based on standard pulse signals sent by the POTS board.

Polarity Reversal Accounting

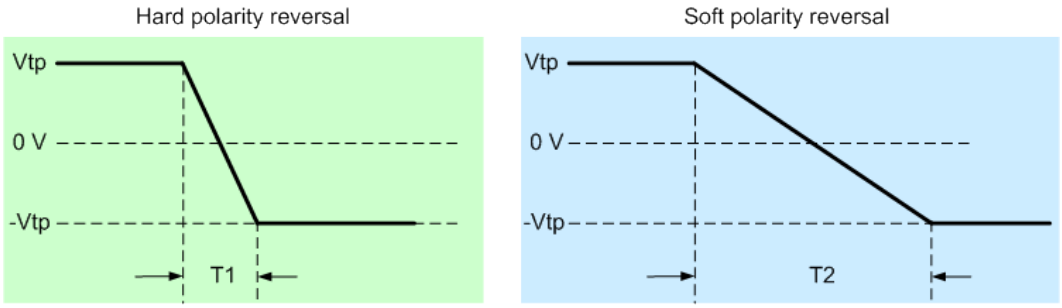
Polarity reversal refers to the voltage reversal between wires. For example, if the voltage between A and B wires is Vtp, the reversed voltage is -Vtp. After detecting a polarity reversal signal, the pay phone starts to account. The pay phone stops accounting when the voltage between A and B wires changes from -Vtp to Vtp, as shown in Figure 23-17. The polarity reversal is classified into hard polarity reversal and soft polarity reversal.

- Hard polarity reversal: also called quick polarity reversal. Specifically, the voltage is reversed in a short period of time, shorter than 3 ms in general. The hard polarity reversal brings great interference on lines.
- Soft polarity reversal: also called slow polarity reversal. Specifically, the voltage is reversed in a long period of time, longer than 80 ms in general. The soft polarity reversal brings small interference on lines.

NOTE

Some terminals are faulty if the polarity reversal time is long. In this situation, the hard polarity reversal must be used, although it brings great interference on lines.

Figure 23-17 Hard and soft polarity reversal



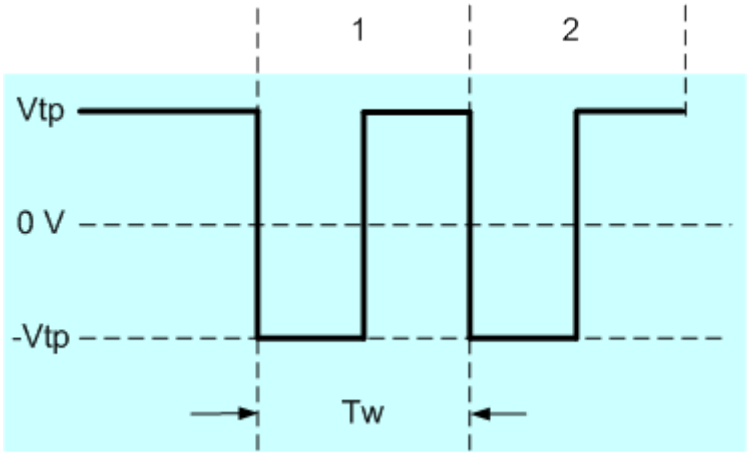
12/16KC Accounting

When detecting that both the calling and called parties enter a session, the softswitch requests the AG to send 12 kHz or 16 kHz high frequency pulse signals to the pay phone. The pay phone performs the accounting once after detecting such a pulse signal. For example, if the carrier charges subscriber 1 cent for one-minute call, the softswitch sends such a pulse signal at the interval of one minute. If the pay phone receives three such pulse signals after the call ends, the subscriber is charged 3 cents. After detecting that the call ends, the softswitch requests the AG to stop sending such pulse signals to the AG.

Polarity Reversal Pulse Accounting

The polarity reversal pulse accounting rule is the same as the polarity reversal accounting rule. The difference lies in that the POTS board sends standard pulse signals to pay phones. As shown in Figure 23-18, the pulse width is Tw and there are a total of two pulse signals.

Figure 23-18 Polarity reversal pulse accounting



23.4.6 Hookflash

Definition

When a phone is in the offhook state, the terminal generates an onhook signal for a period.

Action

To generate hookflash, you can quickly press the hookflash button or **R** (generally) on an ordinary phone.

Application

During a call, the user wants to start some new services. For example, the call forwarding service: user A calls user B. During the call, user B presses the hookflash button and hears the special dial tone. User B dials the number of user C and communicates with user C. User B hangs up the phone, and then user A communicates with user C.

Hookflash Signal

Hookflash is short-time phone onhook actually. However, onhook signals last only a short period of time and therefore hookflash cannot be determined as phone onhook. Both upper and lower thresholds for hookflash are set in the system. When the last duration of an onhook signal is within the range between the upper and lower thresholds, the onhook signal is determined as the hookflash signal.

Upper and lower thresholds for hookflash are defined differently in countries. For example, in China, the upper threshold is 350 ms and lower threshold is 100 ms. This indicates that if an onhook signal lasts 100-350 ms, it is determined as a hookflash signal.

23.4.7 Dual Tone Multi Frequency

Introduction

DTMF means that the tones of two frequencies are overlaid to represent a number, as shown in Table 23-6.

Table 23-6 Mapping between frequencies and numbers

Unit: Hz	1209	1336	1477	1633
697	1	2	3	A
770	4	5	6	B
852	7	8	9	C
941	*	0	#	D

When numbers are dialed on the phone, the dialed numbers are converted into the dual-frequency overlay tones. The DSP detects the dialed numbers by checking the DTMF.

The supported DTMF-specific functions are as follows:

- DTMF erasure: After the DSP detects DTMF signals, it erases the DTMF signals from the RTP media stream.
- DTMF transparent transmission: After the DSP detects DTMF signals, it retains the DTMF signals in the RTP media stream.
- DTMF RFC2833 transmission: After the DSP detects DTMF signals, it erases the DTMF signals from the RTP media stream and sends the DTMF information in RFC2833 transmission mode.

Specifications

- Detection and sending of the DTMF signals is supported.
- Configuration of DTMF-specific functions (device-based) is supported.

Reference Standards and Protocols

ITU-T Q.24

23.4.8 Calling Indication

In a voice call, common calling indications (CINDs) are as follows:

- Call attempt per second (CAPS): Used to measure the volume of concurrent calls.
- Busy hour call attempts (BHCA): Used to measure the system capability of processing calls. BHCA = CAPS/3600.
- Erlang (ERL): Used to measure the traffic. Erl = Calls per hour x Average call hold duration (unit: s)/3600.
- Call loss count: Indicates failed calls.
- Convergence ratio: Used to measure the ratio of system trunk capacity to user capacity.

23.5 SIP Voice Feature

This topic first describes the SIP protocol, and then describes in detail the principle of the SIP protocol.

23.5.1 What Is the SIP Protocol

Definition

Session Initiation Protocol (SIP), defined in RFC 3261, is used for setting up, modifying, and terminating sessions with one or more participants. The session can be a multimedia meeting, distance learning, or Internet telephony. SIP can be used for initiating sessions or inviting a member to join a session that has been set up otherwise. SIP transparently supports the mapping of names and the redirecting service, which facilitates the implementation of ISDN service, intelligent network, and personal mobile service. Once the session is set up, media streams are simply transmitted at the bearer layer through the Real-time Transport Protocol (RTP).

SIP is a text-based protocol put forth by IETF for IP phone/multimedia conferencing. It is a light-weight signaling protocol and has the following features:

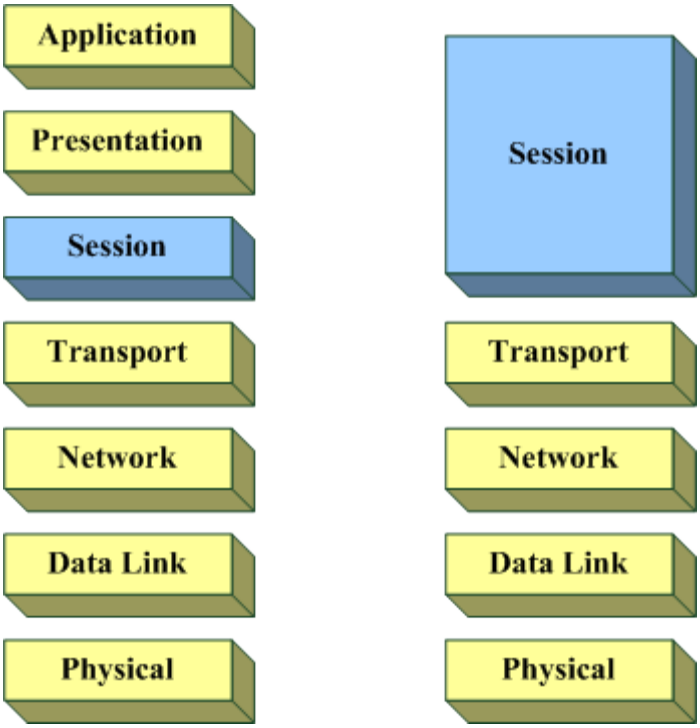
1. Minimum status: One conference call or phone call can contain one or multiple requests or transactions. The proxy server can work in the stateless mode.
2. Irrelevance with lower layer protocols: SIP has minimum assumption of the lower layer protocols. The lower layer protocols can provide reliable or unreliable services to the SIP protocol layer, which can be packet or byte stream services. On the Internet, the SIP protocol layer can use the UDP or TCP protocol, and UDP is preferred. When UDP is not available, TCP is used.
3. Text-based: SIP adopts the text-based UTF-8 coding format and uses the ISO 10646 character set, which makes it easy to realize programming languages such as Java. This feature brings about merits such as easy commissioning, flexibility, and extensibility. The length of message, however, may also increase. For this reason, the message format is particularly designed so that the SIP messages are easy to parse.
4. Robustness: The robustness of SIP is demonstrated in several facets. For example, the proxy server does not need to maintain the call status, subsequent requests and re-transmission can adopt different routes, and the response message is transmitted in the self-routing mode.
5. Extensibility: The extensibility of SIP is demonstrated in several ways. Unidentifiable header fields can be ignored, the user can specify the message content that the SIP server must understand, new header fields can be introduced easily, and status codes are encoded in the layered coding mode.
6. Readiness to support IN services: Working with the end system, SIP and other call control extended protocols can support most services in Capability Set 1 and Capability Set 2 of ITU-T.

Position of the SIP Protocol on the Network

The SIP protocol is a signaling control protocol at the application layer. In the five-layer TCP/IP model, SIP is an application layer protocol. In the seven-layer OSI model, SIP is a session layer protocol. Figure 23-19 shows the position of the SIP protocol on the network.

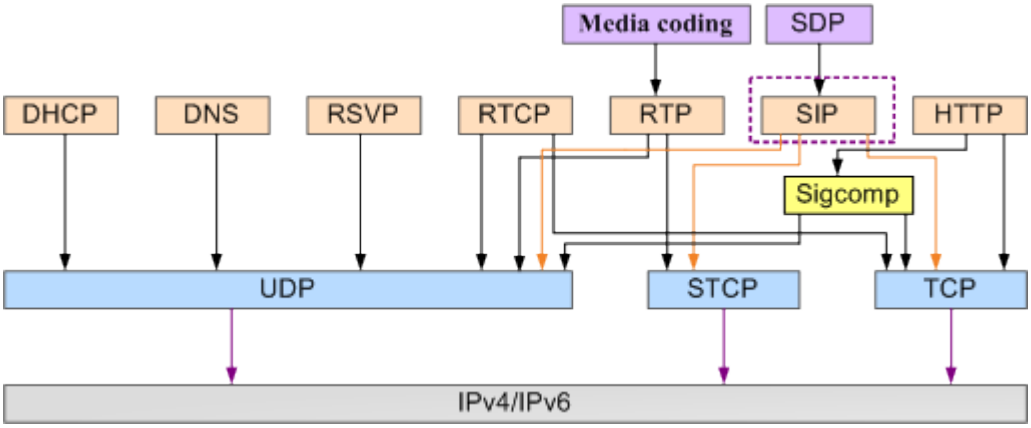
The SIP protocol is independent from transmission protocols but is carried on different transmission protocols, such as, UDP, TCP, TLS, and SCTP. SIP is always carried on the UDP for its efficient transmission.

Figure 23-19 Position of the SIP protocol on the network



The SIP protocol must work together with other protocols to complete multimedia calls. Figure 23-20 shows positions of the SIP and other protocols.

Figure 23-20 Positions of the SIP and other protocols



The SIP protocol works with the Real-time Transmit Protocol (RTP), Real-Time Transport Control Protocol (RTCP), domain name server (DNS), Resource ReServation Protocol (RSVP), and Session Description Protocol (SDP) to complete multimedia calls.

- RTP: A protocol defined by RFC 3550 for transmitting E2E real-time data. It provides the following functions for a series of E2E real-time data transmission services: payload type identification, sequence number arranging, timestamp, and transmission monitoring.
- RTCP: A protocol that controls transmission of real-time media streams.

- RSVP: A protocol that preserves network resources.
- SDP: A text-based application layer protocol for describing multimedia sessions.
- Sigcomp: A mechanism defined by RFC 3320 and used by application layer protocols to compress messages before they are sent to the network.

Advantages of the SIP Protocol

SIP will revolutionize the mode of communication service provisioning and the users' habit of communication consumption. An innovating communication mode integrating video phone service, messaging, Web service, e-mail, synchronous browsing, and conference call will be introduced to the telecommunication industry. Adopting SIP as the control layer protocol has the following advantages:

1. Based on an open Internet standard, SIP has inherent benefits in the integration and interoperability of voice and data services. SIP can implement across-media and across-device call control, and supports various media formats. SIP also supports dynamic adding and deleting of media streams, which make it easier to support richer service features.
2. SIP is intelligently extensible to the service and terminal side, reducing the network load and facilitating the provisioning of service.
3. SIP supports mobile functions at the application layer, including the dynamic registering mechanism, location management mechanism, and redirecting mechanism.
4. SIP supports features such as presence, fork, and subscription, which facilitates development of new services.
5. As a simple protocol, SIP has generally acknowledged extensibility.

23.5.2 Mechanism of the SIP Protocol

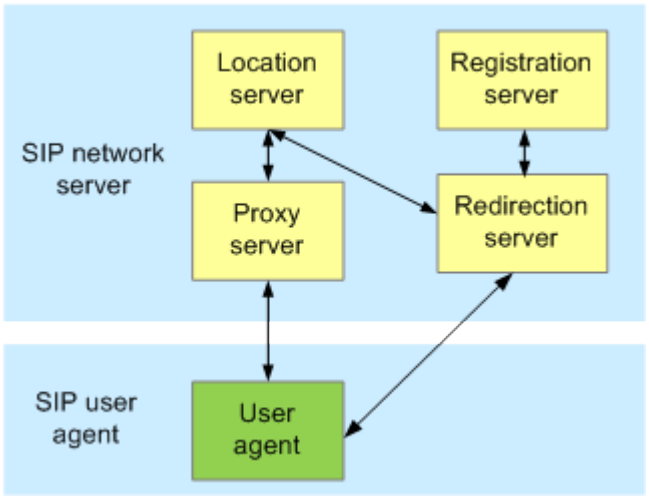
This section describes the SIP protocol that involves network entities, SIP URI, SIP messages, and SIP media negotiation mechanism. Learning of this chapter enables you to have a deep understanding of the SIP protocol.

SIP Network Entities and Application

SIP Network Entities

Various logical entities exist on the SIP network to play different roles. Major SIP entities are SIP user agent and SIP network server, as shown in Figure 23-21.

Figure 23-21 SIP network entities



User agent (UA)

The UA sends or receives SIP requests and processes these requests. The UA is logically classified into user agent client (UAC) and user agent server (UAS). The UAC sends requests to UAS, and the UAS responds to these requests. The responses can be acceptance, rejection, or redirection.

NOTE
A Huawei AG functions as a UA.

Proxy server

The proxy server, a logical network entity, forwards requests or responses on behalf of a client. The proxy server can also function as a server. The proxy server provides the following functions: routing, authentication, accounting control, call control, and service providing. The proxy server attempts to forward requests to multiple addresses in multiple modes, such as branching, cycling, and recursively querying.

Registration server

The registration server receives registration requests and saves address mapping contained in the registration requests to the database for the usage by subsequent call processing and subscriber's home address locating.

Redirection server

The redirection server responds to received requests with one or multiple new addresses. Then, the client simply sends requests to these addresses. The redirection server does not receive or reject calls. It mainly completes the routing function and can support the mobility of SIP terminals together with the registration process.

Location server

The location server provides locating functions. The location server obtains the possible called party's address for the redirection server and proxy server and provides a list of mapping between recorded addresses and contact addresses.

In the actual establishment of a SIP application system, a SIP server must cooperate with other background applications to provide a manageable carrier network. For example, the SIP

server needs to communicate with the Remote Authentication Dial In User Service (RADIUS) server for terminal authentication.

NOTE

The IMS network or softswitch functions as the redirection server, proxy server, and registration server, and the DNS server functions as the location server.

Basic SIP Functions

SIP provides the following basic functions:

- User location: determines terminals used for communication.
- User capabilities: determines the communication media and parameters used by the media.
- User availability: determines the willingness of the called party to join in the communication.
- Call setup: establishes a call between calling and called parties.
- Call handling: transfers or terminates calls.

SIP AG on the IMS Network

Figure 23-22 shows the position of the SIP AG on the IMS network. NEs that have close relationship with the AG include call session control function (CSCF), proxy-call session control function (P-CSCF), interrogating-call session control function (I-CSCF), serving-call session control function (S-CSCF), home subscriber server (HSS), subscription locator function (SLF), media resource server (MRS), and application server (AS).

- CSCF: The call control center of the IMS system. The CSCF dispatches multiple real-time services on the IP transmission platform and provides central routing engine, policy management, and policy implementation.
- P-CSCF: The initial contact point between subscribers and the IMS. The P-CSCF routes terminal requests to a correct I-CSCF or S-CSCF and generates CDRs for roaming subscribers. It provides SIP compression on the Gm interface and integrity protection.
- I-CSCF: During the IMS terminal registration, the I-CSCF assigns an S-CSCF for processing subscriber services and locates the S-CSCF that the called party registers with.
- S-CSCF: The S-CSCF provides registration, authentication, service triggering and control, and session routing functions for IMS subscribers.
- HSS: The HSS functions as a database to store information, such as subscriber numbers.
- SLF: When multiple HSS devices exist in the domain, the SLF selects an HSS for storing subscriber data.
- MRS: It is classified into multimedia resource function controller (MRFC) and multimedia resource function processor (MRFP). The MRS plays tones and announcements, processes conference media stream (audio mixing) and DTMF digits, and converts codecs. In some situations, the MRS and AS functions can be played by one device.
- AS: It triggers services.

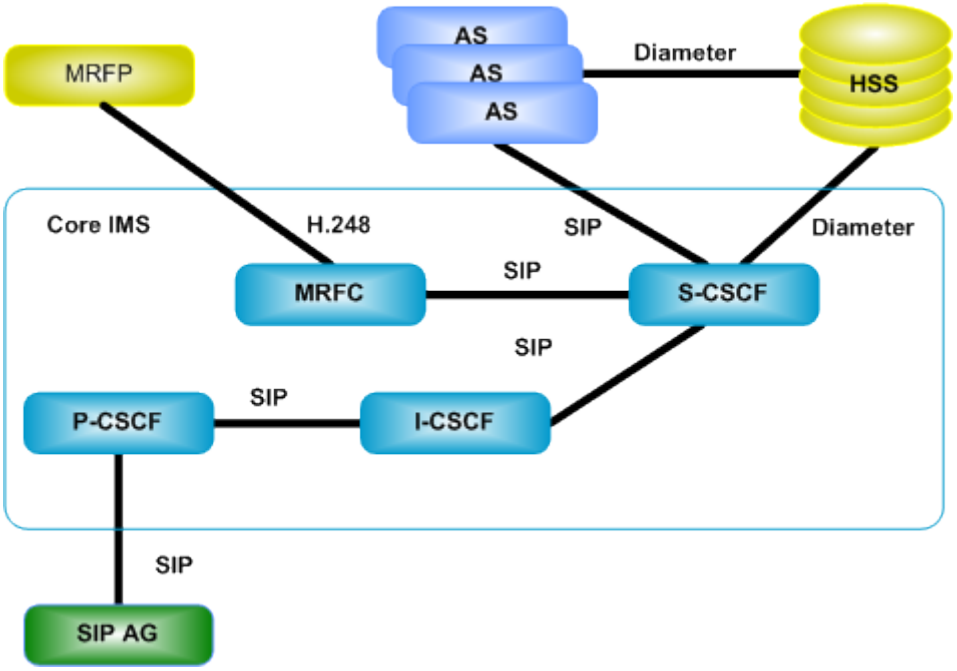
Figure 23-22 shows voice service providing using the SIP protocol on the IMS network.

- The SIP AG accesses the IMS network through the P-CSCF. The P-CSCF forwards SIP messages (including registration, multimedia session, and IM/Presence messages) to the

homed S-CSCF (based on registration information) or I-CSCF (based on the home domain name carried in the SIP message).

- The I-CSCF assigns an S-CSCF for processing subscriber information based on subscriber registration information and CSCF capability information.
- The S-CSCF receives registration requests sent by IMS subscribers and forwarded by the P-CSCF, cooperates with the HSS to authenticate subscribers, and provides routing functions for calling and called parties. The IMS communicates with the AS and MRFC under the control of the S-CSCF.

Figure 23-22 Simplified IMS networking



SIP URI

The SIP protocol uses the uniform resource identifier (URI) to identify terminal users. RFC 2369 defines URI rules and syntax.

SIP URI includes SIP URI and TEL URI, either of which can uniquely identify a SIP user. The SIP URI configured on the MA5600T&MA5603T&MA5608T and the IMS for a SIP user must be the same.

SIP URI

SIP URI is used in SIP messages, indicating the initiator of a request (From), the current destination address (Request-URI), the final receiver (To), and the address after redirection (Contact). SIP URI can also be embedded into the Web page or other hyper links to indicate that a certain user or service can be accessed through SIP.

Generally, the SIP URI is in the following format:
sip:user:password@host:port;URI-parameters?headers

Table 23-7 SIP URI format description

Item	Description
SIP	Indicates that the SIP protocol is used to communicate with the peer end.
user	Indicates the user name, which can consist of any characters. In general, the user name can be an e-mail address or a telephone number.
password	Indicates the password, which can be presented in the SIP URI. However, password presentation in the SIP URI is not recommended, because it poses security risks.
host	Indicates the host name. It can be the host domain name or an IPv4 address.
port	Indicates the ID of the port to which requests are sent. The default value is **5060**, which is the ID of the public SIP port.
URI-parameters	Indicates URI parameters. • transport-param: specifies the protocol used in the transmission layer, such as, UDP or TCP. • user-param: identifies whether the user name is a telephone number or a common user name. • method-param: specifies the method that is used. • ttl-param: indicates the time-to-live (TTL) value of UDP multicast packets. The parameter is used only when the transmission protocol is UDP and the server address is a multicast address. • maddr-param: indicates the address of the server that communicates with the user. The parameter overwrites any address derived from the host field. Generally, the parameter is a multicast address. **NOTE** Parameters transport-param, method-param, ttl-param, and maddr-param are all URL parameters. They are used only in a redirected address, that is, the Contact header field.
headers	Is contained in requests. The header field can be specified using a question mark (?) in a SIP request. For example, sip:alice@example.huawei.com?priority=urgent

Table 23-8 shows examples of the SIP URI.

Table 23-8 SIP URI examples

Example	Description
sip:55500200@191.169.1.112;	**55500200** indicates the user name, and **191.169.1.112** indicates the IP address of the gateway for IP calls.
sip:55500200@ 191.169.1.112 :5061; User=phone;	**55500200** indicates the user name, **191.169.1.112** indicates the IP address of the host, and **5061** indicates the port ID of the host. **User=phone** indicates that the user name is a telephone number.
sips:1234@10.110.25.239	**sips** indicates the secure SIP URI, that is, the security-based TLS protocol is used on the transmission layer. **1234** indicates the user name, and **10.110.25.239** indicates the IP address of the

Example	Description
	gateway for IP calls.

TEL URI

TEL URI identifies a telephone number that occupies resources. The telephone number can be a global number or a local number. The global number must comply with the E164 coding standard and start with a plus sign (+). The local number must comply with local private numbering plan. Format:

tel:+86-755-6544487

tel:45687;phonecontext=example.com

tel:45687;phonecontext=+86-755-65

SIP Message

Format

The SIP message is encoded in the text format, each line ending with CR or LF. The SIP message has two types, the request message and the response message. The general message format is as follows:

SIP message = Start line

*Message header field

Empty line (CRLF)

[Message body]

A SIP message consists of a start-line, one or more header fields, and a message body. The request and response messages are the same in the format and only differ in the start-line. The request message has a request-line as the start-line and the response message has a status-line as the start-line.

Request Message

Request messages are sent from the client to the server. SIP request messages include INVITE, ACK, OPTIONS, BYE, CANCEL, REGISTER, PRACK, and UPDATE. Table 23-9 describes these SIP request messages.

Table 23-9 SIP request messages

Request Message	Function
INVITE	Invites a user to join a call

Request Message	Function
ACK	Acknowledges the response message of the request
OPTIONS	Requests for querying capability information
BYE	Releases an established call
CANCEL	Releases an unestablished call
REGISTER	Registers the user location information on the SIP network server
PRACK	Acknowledges a reliable provisional response message
UPDATE	Updates the session

start-line

The start-line consists of Method, Request-URI, and SIP-Version.

- Method: determines the type and purpose of a request message. Keywords are request messages in Table 23-9.
- Request-URI: identifies the user or server address used by the request.
- SIP Version: indicates the SIP version contained in a request or response. This parameter is case insensitive. In general, values are always in upper cases.

Header field

For details, see Message Header.

Message body

For details, see Message Body.

Response Message

The SIP response message is used for responding to the SIP request message, indicating whether the call is successful or fails. Different from request messages, the start-line of response messages is also called status-line, which consists of SIP-Version, Status-Code, and Reason-Phrase.

- SIP-Version: indicates the used SIP version.
- Status-Code: identifies the response message type. The status-code is a 3-digit integer. The first digit of the status-code defines the response type. The other two digits provide detailed descriptions about the response. Table 23-10 describes response messages.
- Reason-Phrase: provides descriptions about the status code. This field is optional.

Table 23-10 SIP response messages

Status Code	Meaning	Function
1XX	Provisional	The request has been received and is being processed.

Status Code	Meaning	Function
2XX	Success	The action was successfully received, understood, and accepted.
3XX	Redirection	Further action needs to be taken in order to complete the request.
4XX	Client Error	The request contains bad syntax or cannot be fulfilled at this server.
5XX	Server Error	The server failed to fulfill an apparently valid request
6XX	Global Failure	The request cannot be fulfilled at any server.

NOTE
Except 1XX responses, other responses are final responses and can terminate requests.

SIP requires that the application must understand the first integer of the response status code, and allows the application not to process the last two integers of the status code.

For example, SIP/2.0 200 OK

- SIP/2.0 indicates that the SIP version is 2.0.
- 200 is the status code, indicating a successful response.
- OK is the cause code and is the further explanation about 200.

Message Header

The SIP message header consists of SIP header fields to complete information transfer and parameter negotiation for SIP sessions. RFC 3261 defined various SIP header fields. This section only introduces five mandatory header fields in a SIP message.

Table 23-11 Mandatory SIP header fields

Header Field	Function	Common Format	Description
Call-ID	Globally identifies a session.	Local flag@host	-
From	Indicates the initiator of the request. The server copies this field from the request message to the response message.	Displayed name<SIP-URI>;tag=XXXX	The tag, a hexadecimal character string, is used to identify two subscribers who share a SIP address to initiate calls using the same call ID. The tag value must be unique globally. A subscriber must have the same call ID and tag value for a call.
To	Indicates the recipient of the request. It has the same format as From. To and From differs only in the first keyword.	Displayed name<SIP-URI>;tag=XXXX	

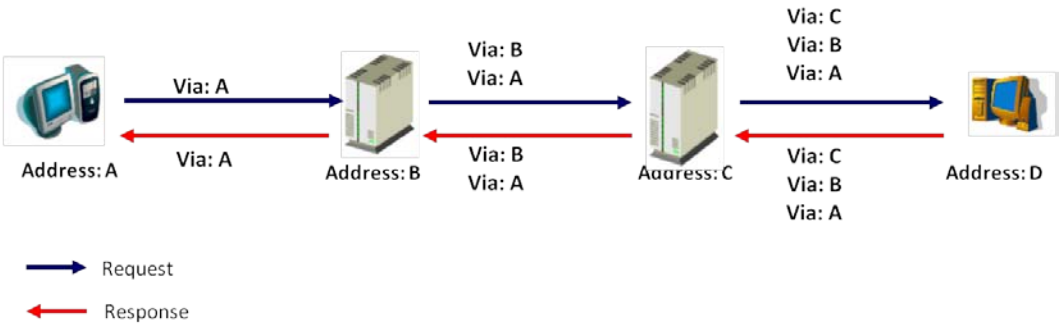
Header Field	Function	Common Format	Description
CSeq	Indicates the sequence number of a request. The client adds this field to each request. The server copies the CSeq value in the request to the response. CSeq is used to determine mapping between responses and requests.	sequence number message name	-
Via	Indicates the path that the request passes. This header field can prevent loop from happening during the request transmission and ensure that the response and request are transmitted in the same path to meet specified requirements.	transmit protocol, sender; hidden parameter, TTL parameter, multicast address parameter, receiver flag, branch parameter	For details, see Via processing.

Via processing

Figure 23-23 shows the Via processing.

- For a request: When transmitting a request, the SIP entity adds its address on the most outer of the Via set of the request. Therefore, when the request reaches the destination, a Via header field set similar to the stack is formed in the request.
- For a response: The destination entity copies the Via address in the request to the response. When receiving the response, the proxy server checks whether the Via at the most outer of the Via set is the proxy server's address. If yes, the proxy server deletes the Via, checks the next Via address, and sends the response to the next Via address. If the next Via address does not exist, this response is terminated on this proxy server.

Figure 23-23 Processing of the Via header field



Message Body

The message body is used to negotiate information and parameters during the call establishment. In addition, the message body also transfers authentication information. In general, SIP messages are in Session Description Protocol (SDP) format.

SIP Media Negotiation Mechanism

Session Initiation Protocol (SIP) cooperates with the Session Description Protocol (SDP) to complete media negotiation. Negotiated media includes the IP address, port ID, codec, and media channel parameters.

SDP

SDP, a text-based control protocol at the application layer, is used for negotiating media, including media type and codec solution during session establishment. For the SDP message format, see RFC 2327. Descriptions of common SDP lines are as follows.

SDP lines	Description
v line	Indicates the SDP protocol version.
o line	Provides the session initiator (user name and host address), session flag, and session version number.
s line	Indicates the session name. Each session has a unique session name in the session description.
c line	Indicates the linked address. In general, the linked address is the IP address of an AG or a SIP terminal controlled by the IP multimedia subsystem (IMS).
t line	Indicates the start and stop times for a session. The t line is used if a session is active at multiple irregularly spaced times.
m line	Indicates media description. A session may not have media description or have multiple media descriptions. Media can be audio, video, application (such as whiteboard information), data, and control.
a line	Indicates media attribute. A session may not have media attribute or have multiple media attributes. The media attribute can be: • media direction: sendonly, recvonly, inactive, or sendrecv • ptme (packetization time) • bandwidth • codec format

Media Negotiation Process

The SIP protocol implements media negotiation based on a simple offer/answer model. In this model, the call initiator informs the call receiver of all supported media formats. The call receiver selects one or multiple media formats to respond the call initiator. Then, media streams are transmitted using negotiated media formats.

Through the negotiation, both the call initiator and call receiver obtain peer media attributes so that they can communicate using correct media channels.

Offer/answer processes cannot be overlapped. Each media negotiation must be on the basis of the previous negotiation. Media streams can be added to (not deleted from) a new SDP offer. Nevertheless, you can set the port number of the media streams to 0 to change the number of media streams that are used.

Media Negotiation Example

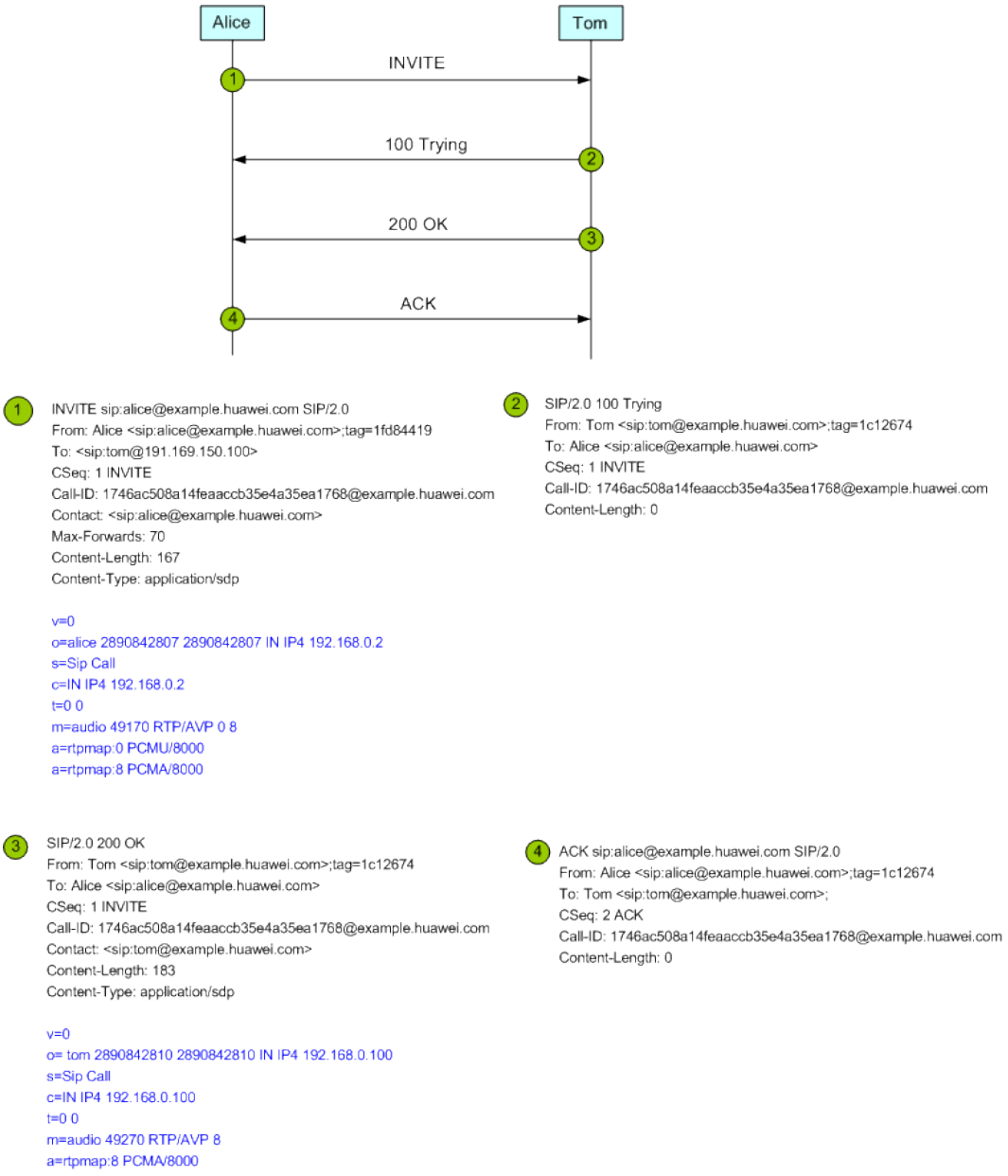
Figure 23-24 shows a SIP media negotiation example.

1. The Alice's SIP phone sends an INVITE message to the Tom's SIP phone to establish a session. The INVITE message contains Alice's SDP information (in blue). The SDP information contains IP address 192.168.0.2, port number 4917, and supported codecs PCMU and PCMA.
2. Tom's SIP phone replies with a 100 response, indicating that the invitation is being processed.
3. Tom's SIP phone sends a 200 response to Alice's SIP phone. The 200 response indicates that the invitation has been successfully processed and carries Tom's SDP information (in blue). The SDP information contains IP address 192.168.0.100, port ID 49270, and supported codec PCMA.
4. After receiving the 200 response, Alice's SIP phone sends an ACK message to Tom's SIP phone to confirm the receiving of the 200 response. Then, Alice and Tom can exchange RTP media packets using the negotiated IP address, port, and codec to communicate with each other.

NOTE

After the session is established, if the Tom's or Alice's SIP phone wants to modify parameters, they can send a Re-INVITE message to initiate another negotiation. The processing rule is the same as that for processing the INVITE message.

Figure 23-24 SIP media negotiation example



23.5.3 SIP Services and Basic Service Flows

This chapter describes the subscriber registration and authentication flow, subscription flow, and basic call flow.

User Registration and Authentication Flows

Before initiating a call, a SIP user must register with the home network to map the domain name to an IP address. The registration is of two types: the registration not requiring authentication and the registration requiring authentication. After the system is powered on or after the user is added, the user registration flow is started.

Registration Not Requiring Authentication

As shown in Figure 23-25, the SIP AG sends the REGISTER request message to the IMS for each user. The message contains information such as the user ID. After receiving the REGISTER request message, the IMS checks whether the user is already configured on the IMS. If the user is already configured, the IMS responds to the SIP AG with the RESPONSE 200 message.

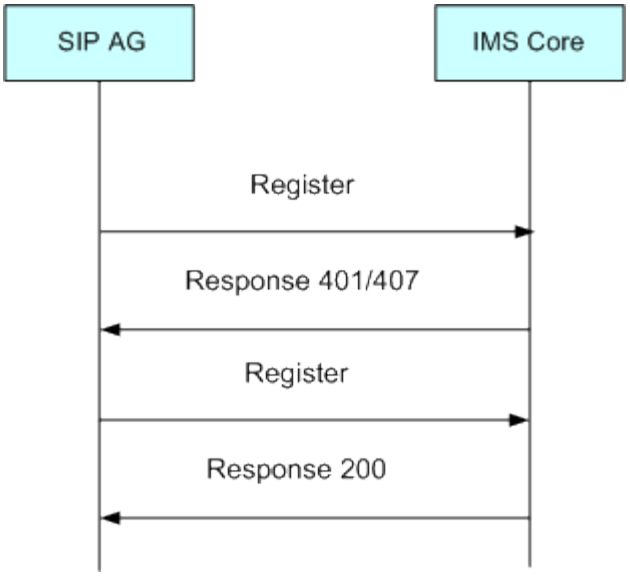
Figure 23-25 Flowchart of the registration not requiring authentication

Registration Requiring Authentication

As shown in Figure 23-26, the SIP AG sends the REGISTER request message to the IMS for each user. The message contains information such as the user ID.

The IMS responds with the RESPONSE 401/407 message, the message containing information such as the key and the encryption mode. The SIP AG encrypts the corresponding user name and password, generates a new REGISTER request message, and sends the message to the IMS. The IMS decrypts the message and verifies the user name and password. If the user name and password are correct, the IMS responds to the SIP AG with the RESPONSE 200 message.

Figure 23-26 Flowchart of the registration requiring authentication



Registration Modes

Table 23-12 shows the registration modes supported by SIP AGs.

Table 23-12 Registration modes supported by SIP AGs

Registration Mode	Description
Separate account registration	A non-wildcard number is used for registration. Each account is registered separately.
Wildcard number registration	A wildcard number, such as 2878*, is used for registration. After the registration, all numbers with prefix "2878" are successfully registered on the IMS.
Proxy group registration	A batch of accounts are added to a group. Then, an account in the group or a separate group account is used for registration. After the registration, all accounts in the group are successfully registered on the IMS.

NOTE
If a large number of accounts concurrently register with the IMS in separate account registration, mass registration messages degrade IMS performance. The wildcard registration and proxy group registration reduce the number of registration messages to be exchanged between AGs and the IMS.

Deregistration

When a SIP AG attempts to deregister from the IMS, the AG sends a REGISTER request with timeout duration set to 0s to the IMS. The deregistration request can also be initiated by the IMS.

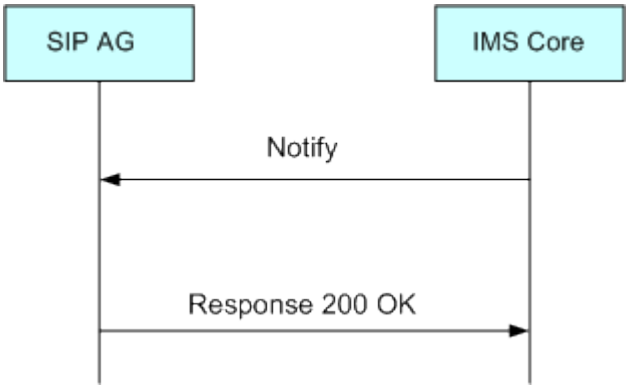
Deregistration Initiated from the AG

Figure 23-27 Deregistration initiated from the AG

As shown in Figure 23-27, the SIP AG initiates deregistration by sending a new REGISTER request with **Expires** set to 0 to the IMS.

Deregistration Initiated from the IMS

Figure 23-28 Deregistration initiated from the IMS



As shown in Figure 23-28, when the IMS requires to clear the registration of a SIP AG, the IMS sends a NOTIFY message carrying the deregistration reason to the AG if the IMS has subscribed to the registration status of this AG. After receiving the NOTIFY message, the SIP AG responds to the IMS with a 200 OK message.

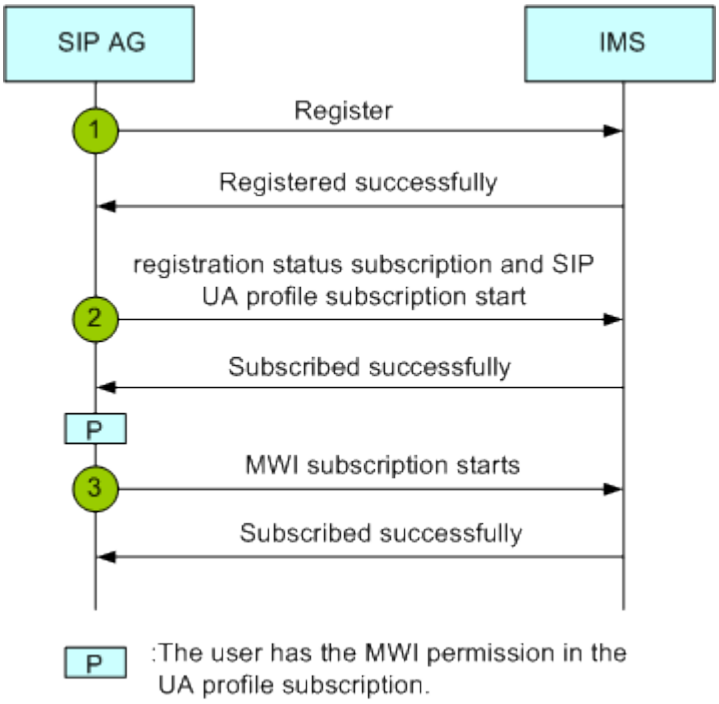
Subscription

Subscription indicates that a SIP access gateway (AG) requests the current status, information change, and service permission of a subscription user from the network side. The SIP AG supports the following subscriptions:

- Registration status subscription
- User agent (UA) profile subscription
- Message waiting indication (MWI) subscription

As shown in Figure 23-29, after a user is successfully registered, registration status subscription and SIP UA profile subscription start. If the user has the MWI permission in the UA profile subscription, an MWI subscription starts.

Figure 23-29 Subscription flow

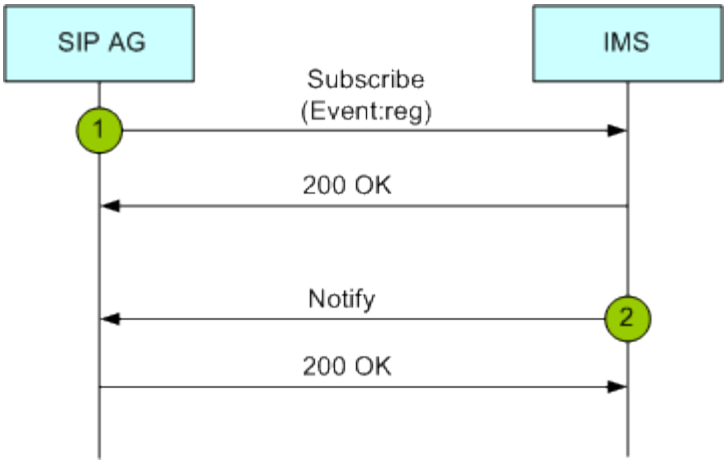


Registration Status Subscription

Registration status subscription is mainly used to obtain the user's registration status and registration status change. Figure 23-30 shows the flow of a registration status subscription:

1. The SIP AG uses the **Subscribe** message to initiate a registration status subscription and uses the **Event** header to identify the subscription type.
2. The IMS uses the **Notify** message to issue the user's registration status to the SIP AG.

Figure 23-30 Flow of a registration status subscription



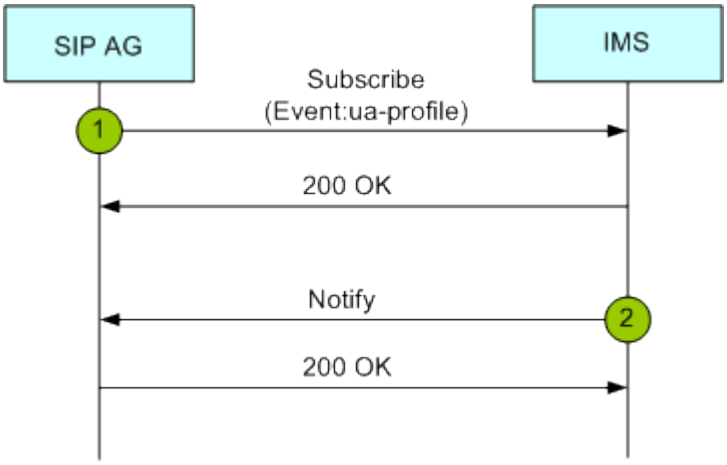
UA Profile Subscription

UA profile subscription is used to obtain the user's service permission and dial tone scheme. For details about service permissions supported by a SIP user, see 23.5.4 SIP Value-added Services. Figure 23-31 shows the flow of a UA profile subscription.

1. The SIP AG uses the **Subscribe** message to initiate a registration status subscription and uses the **Event** header to identify the subscription type.
2. The IMS uses the **Notify** message to issue the user's registration status to the SIP AG.

NOTE
Users' service permissions can be subscribed to the IMS and can be configured at local through the SIP AG. If the SIP AG is not subscribed to the IMS, service permissions configured on the SIP AG take effect.

Figure 23-31 Flow of a UA profile subscription

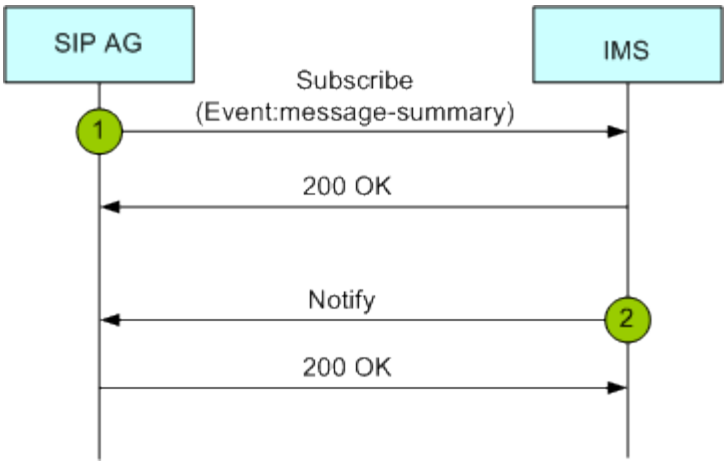


MWI Subscription

MWI subscription is used to obtain the user's messages. Figure 23-32 shows the flow of an MWI subscription.

1. The SIP AG uses the **Subscribe** message to initiate a registration status subscription and uses the **Event** header to identify the subscription type.
2. The IMS uses the **Notify** message to issue the user's registration status to the SIP AG.

Figure 23-32 Flow of an MWI subscription



SIP-based VoIP

Figure 23-33 shows the networking for SIP-based voice over IP (VoIP) calls. SIP-based VoIP services include plain old telephone service (POTS) services, integrated services digital network (ISDN) services, and R2 services.

Figure 23-33 Networking for SIP-based VoIP calls

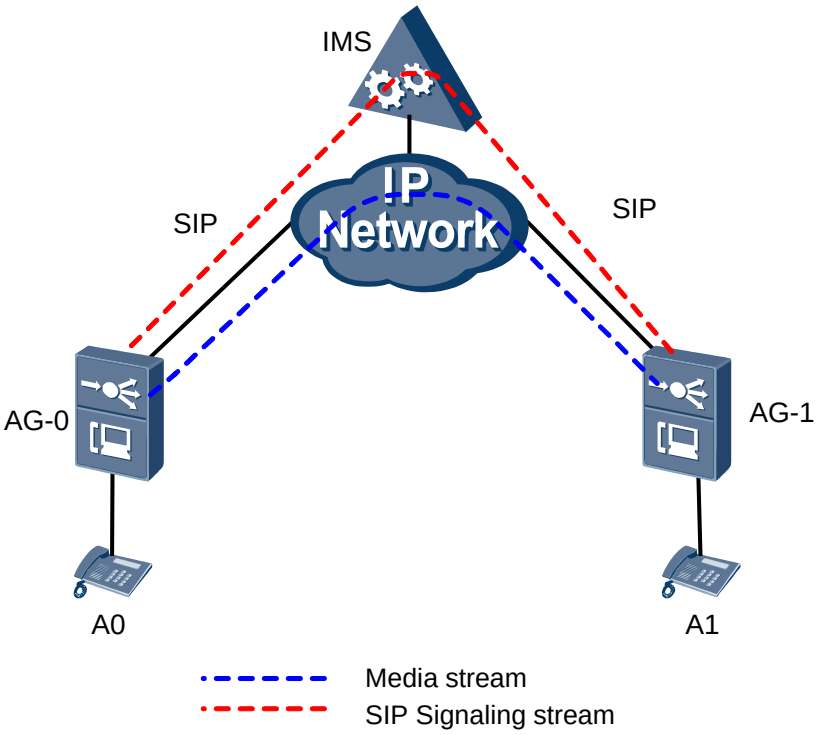
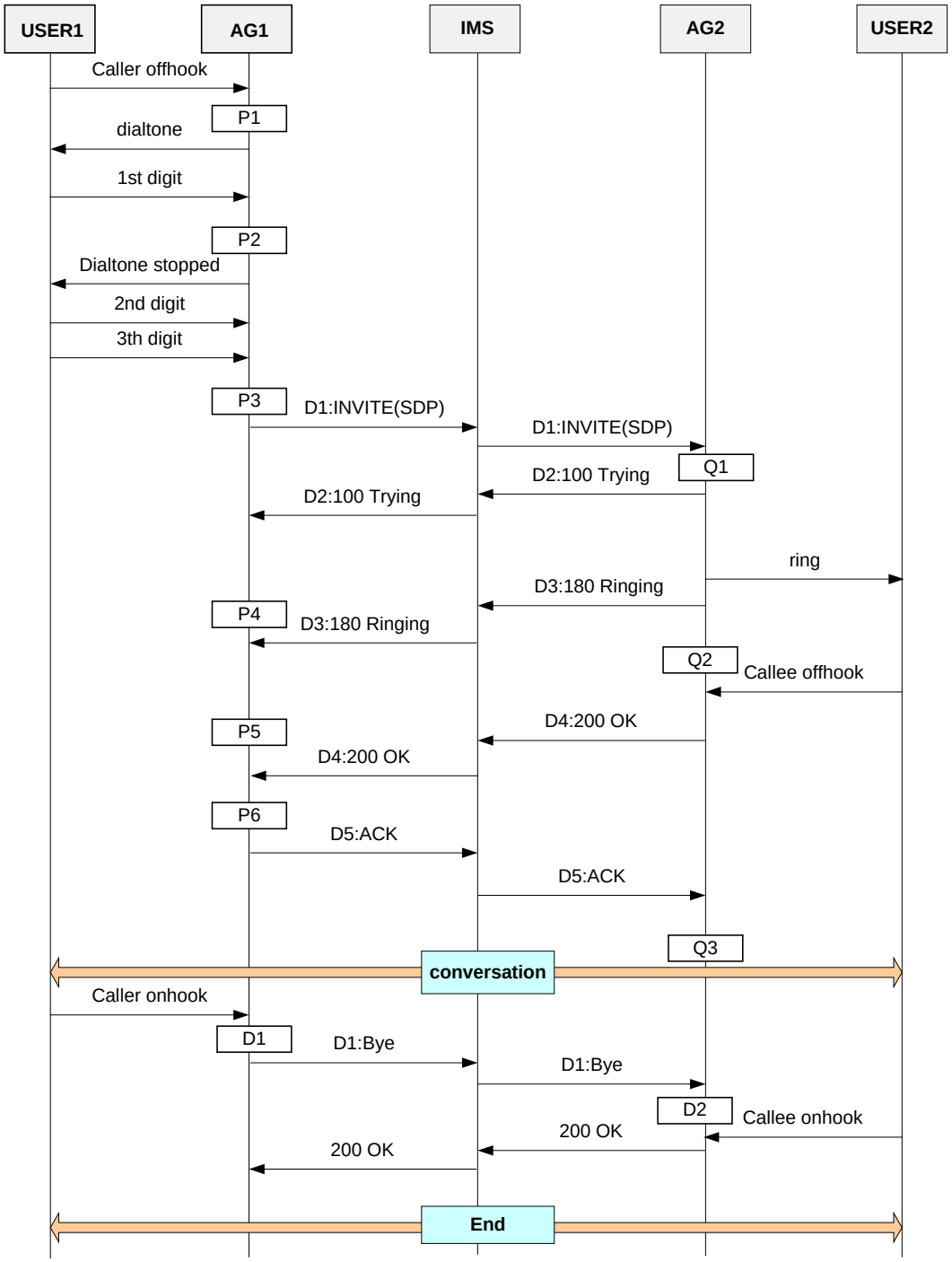


Figure 23-34 shows the SIP-based VoIP call flow. USER1 is the calling party, and USER2 is the called party.

Figure 23-34 SIP-based VoIP call flow



Call flow on the calling side

- P1: AG1 receives the offhook message of USER1 and plays the dial tone to USER1.
- P2: AG1 receives the first dialed digit, stops playing the dial tone, and then starts matching the digit with the digitmaps.

- P3: After receiving N dialed digits and matching the digits with the digitmaps, AG1 finds that the dialed number matches a certain digitmap. Then, AG1 generates the INVITE message and sends the message to the IMS.
- P4: AG1 receives RESPONSE 100 and knows that the peer end receives the INVITE message, so AG1 stops the INVITE message retransmitting flow.
- P5: AG1 receives 180, which indicates that the phone of USER2 is ringing. Then, AG1 plays the ring back tone to USER1.
- P6: AG1 receives 200, which indicates that USER2 answers the phone, so AG1 stops playing the ring back tone to USER1, and changes the stream mode to the bidirectional mode. Then, AG1 constructs an ACK message and sends the message to the IMS.

In the actual processing, when USER1 initiates a call, the IMS determines the situation as follows:

- If USER1 has been configured but not registered with the IMS, the IMS rejects USER1 and responds with 403 to AG1.
- If USER1 is not configured, the IMS rejects USER1 and responds with 404 to AG1.

Call flow on the called side

- Q1: After receiving the INVITE message from the IMS, AG2 replies with a 100 response to the IMS, finds USER2 based on the P-Called-Party-ID, RequestURI, and TO fields (or TEL-URI) contained in the INVITE message, plays the ring tone to USER2, and sends 180 to the IMS, indicating that USER2 is being alerted.
- Q2: After detecting that USER2 picks up the phone, AG2 stops playing the ring tone and sends 200 to the IMS, indicating that USER2 has picked up the phone.
- Q3: After AG2 receives an ACK message, calling and called parties start a session.

In the actual processing, when USER1 initiates a call, AG2 determines the situation as follows:

- If USER2 has been configured but is not registered with the IMS, AG2 replies with 403 to reject the call.
- If USER2 is not configured, AG2 replies with 404 to reject the call.

Call release flow

- D1: After detecting that USER1 hangs up, AG1 sends a BYE message to the IMS and releases DSP resources.
- D2: After receiving the BYE message, AG2 notifies USER2 of the on-hook event. USER2 hangs up, and the call ends.

SIP-Based FoIP

In terms of transmission protocol, the fax service can be classified into transparent transmission and T.38; in terms of switching mode, the fax service can be classified into auto-switching and negotiated-switching. Hence, there are four combinations of the fax mode: auto-switching transparent transmission, auto-switching T.38, negotiated-switching transparent transmission, and negotiated-switching T.38.

The working principle of auto-switching is that the AG detects the fax tone, and then selects the transparent transmission or T.38 mode according to the configuration. In this case, the AG needs not send any signaling to the peer device.

The working principle of negotiated-switching is that the AG detects the fax tone, and according to the configuration sends the peer end the re-INVITE message that contains the negotiation parameters for negotiating the fax mode.

In actual application, fax can also be classified into low-speed fax and high-speed fax in terms of transmission speed. The high-speed fax cannot adopt the T.38 mode. A high-speed fax machine can actually be regarded as a modem. With the speed reduced, a high-speed fax machine can also adopt the T.38 mode.

Flow of the Negotiated-Switching Transparent Transmission Fax

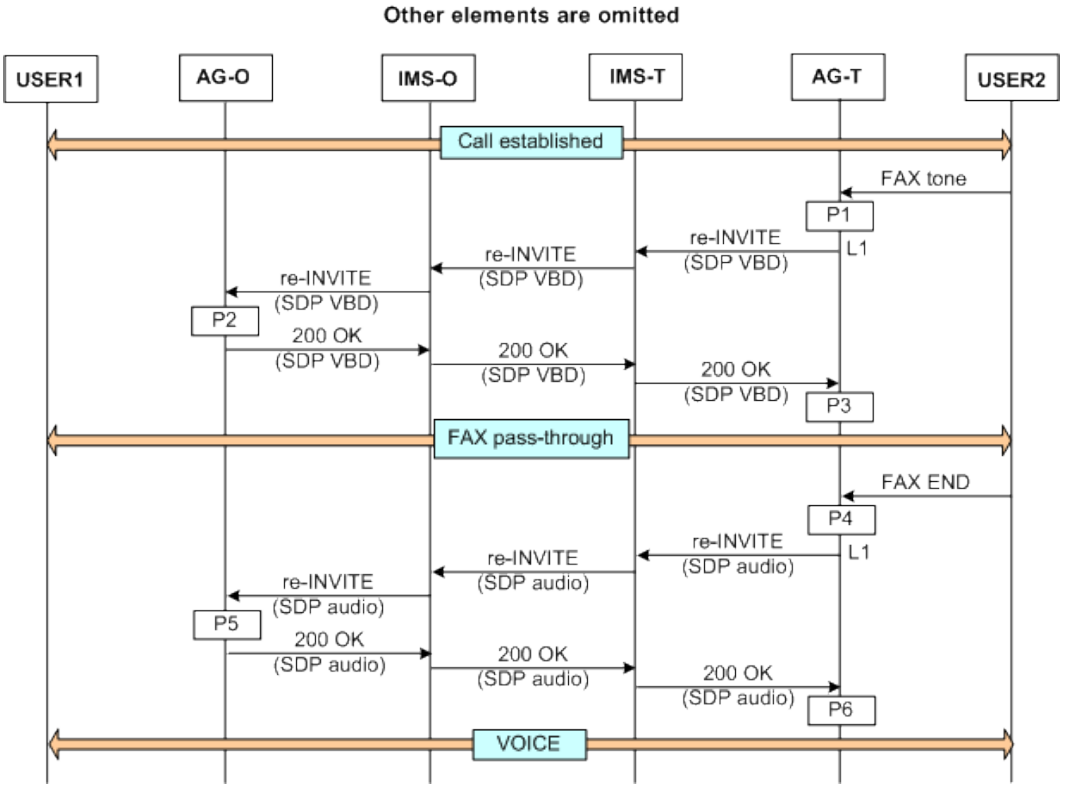
Currently, this fax mode can be presented in three ways.

- Presented as a=fax. This is a G.711 transparent transmission fax mode proposed by China Telecom.
- Presented as a=silenceSupp:off. This is a G.711 transparent transmission fax mode defined in the *draft-IETF-sipping-realtimefax-01.txt*.
- Presented as a=gpmd:99 vbd=yes. This is a VBD mode defined in the ITU-T V.152.

Which method to be applied depends on the parameters configured.

Figure 23-35 shows the fax flow.

Figure 23-35 Flow of the negotiated-switching transparent transmission fax



- P1: AG-T first detects the fax tone, and then sends the re-INVITE message to the AG (AG-O) to which the calling party is connected.
- L1: The SDP message contained in the re-INVITE message has three types. The specific fax mode must be configured on the AGs. The initiator of negotiation uses the **a**

parameter of different values, and the recipient of negotiation needs to be compatible with the three parameter values. This means that when the recipient receives the re-INVITE message, the recipient should be able to complete the negotiation process with the initiator regardless of the **a** parameter value.

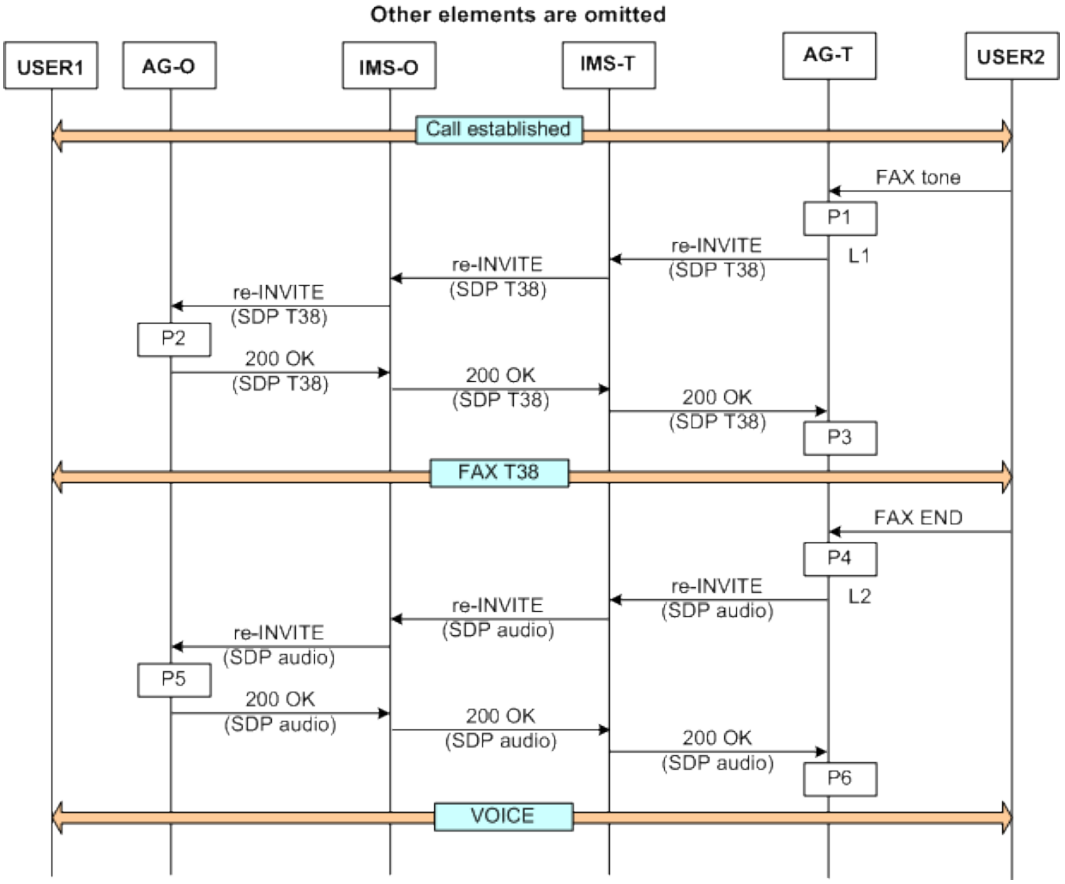
  - The G.711 transparent transmission fax/modem mode defined in the *draft-IETF-sipping-realtimefax-01.txt*.
  - The G.711 transparent transmission fax/modem mode proposed by China Telecom.
  - The VBD mode defined in the ITU-T V.152.

- P2: AG-O receives the re-INVITE message. Then, AG-O generates the 200 OK message and sends the message to AG-T.
- P3: AG-T receives the 200 OK message, and also enables the DSP channel in the fax mode.
- P4: AG-T receives the fax end signal, and sends the re-INVITE message to AG-O.
- L2: The SDP message contained in the re-INVITE message is for setting up a common voice channel.
- P5: AG-O receives the re-INVITE message and switches the DSP channel to the voice mode.
- P6: AG-T receives the 200 OK message, and also switches the DSP channel to the voice mode.

Flow of the Negotiated-Switching T.38 Fax

Figure 23-36 shows the flow of the negotiated-switching T.38 fax.

Figure 23-36 Flow of the negotiated-switching T.38 fax



- P1: AG-T first detects the fax tone, and then sends the re-INVITE message to the AG (AG-O) to which the calling party is connected.
- L1: The SDP message contained in the re-INVITE message carries the T.38 information.
- P2: AG-O receives the re-INVITE message, learns that the peer device requires the T.38 mode, and enables the DSP channel in the T.38 mode. Then, AG-O generates the 200 message and sends the message to AG-T.
- P3: AG-T receives the 200 OK message, and also enables the DSP channel in the T.38 mode.
- P4: AG-T receives the fax end signal, and sends the re-INVITE message to AG-O.
- L2: The SDP message contained in the re-INVITE message is for setting up a common voice channel.
- P5: AG-O receives the re-INVITE message and switches the DSP channel to the voice mode.
- P6: AG-T receives the 200 OK message, and also switches the DSP channel to the voice mode.

NOTE

Figure 23-37 and Figure 23-38 shows the fax flows when the peer device does not support the T.38 mode.

Figure 23-37 Flow of the negotiated-switching T.38 fax when the peer device does not support the T.38 mode (scenario 1)

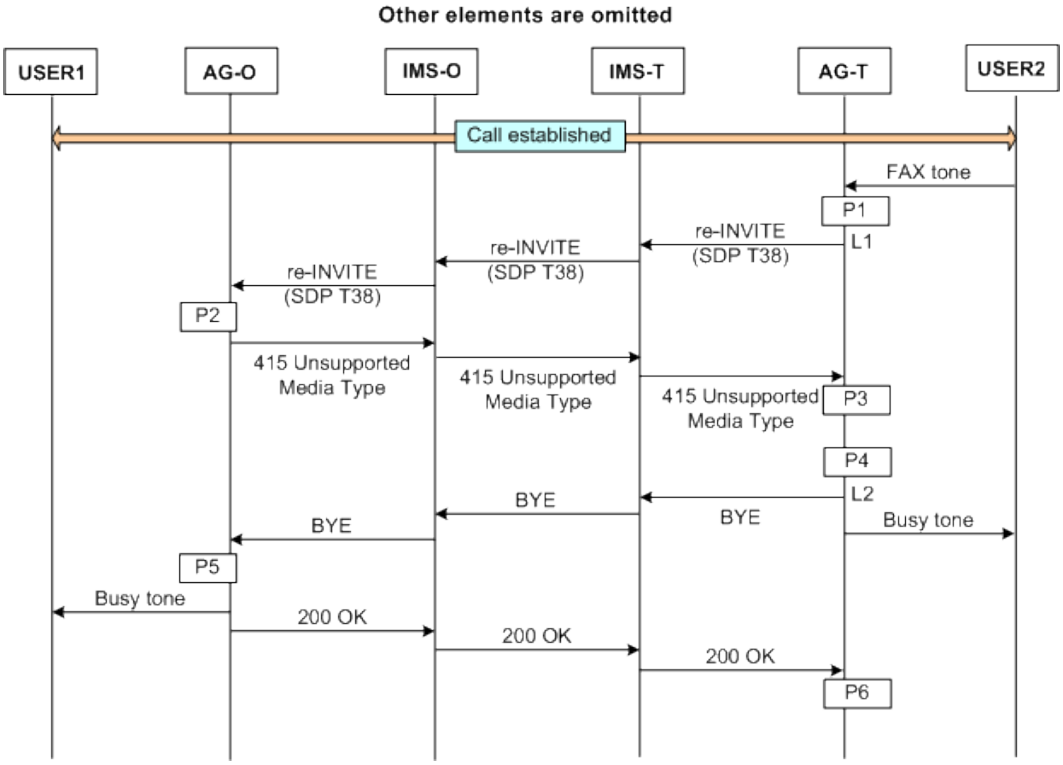
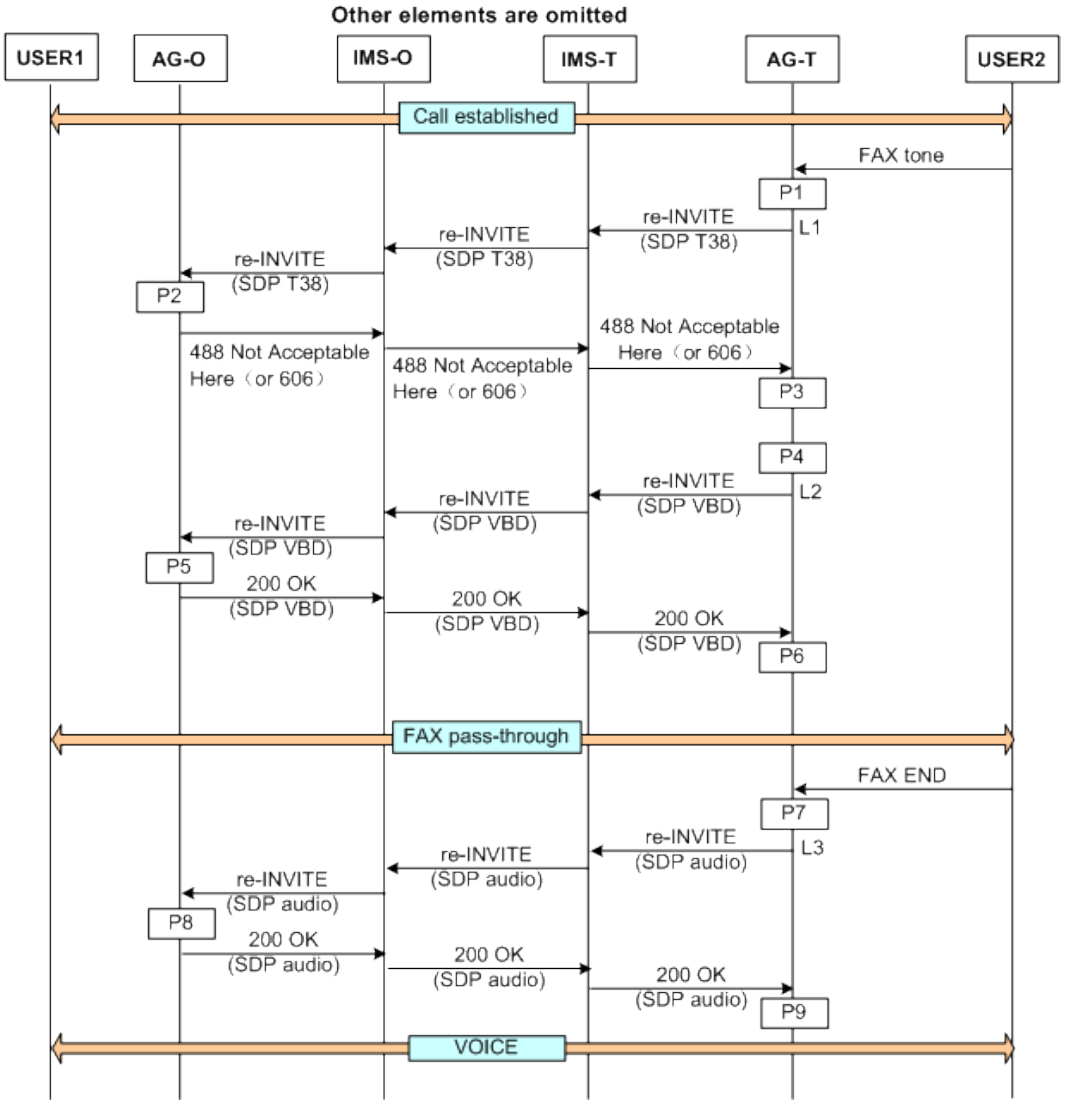


Figure 23-38 Flow of the negotiated-switching T.38 fax when the peer device does not support the T.38 mode (scenario 2)



In scenario 1, if AG-O does not support T.38, it may respond with 415 Unsupported Media Type. After AG-T receives the 415 response, AG-T sends the BYE message and releases the current call. In scenario 2, if AG-O does not support T.38, it responds with 488 Not Acceptable Here or 606 Not Acceptable. After AG-T receives the 488/606 response, AG-T generates another re-INVITE message. The SDP message in this message contains the VBD media type. Thus, the negotiation on the T.38 mode fails, and the transparent transmission mode is adopted.

The MA5600T/MA5603T/MA5608T supports the T.38 mode, and therefore does not respond with the 415/488/606 message in the T.38 negotiation. The MA5600T/MA5603T/MA5608T, however, can process such error codes sent by the peer device.

Flow of the Auto-Switching Transparent Transmission Fax

Generally, the called fax terminal detects the fax tone on the TDM side first, and the calling fax terminal detects the fax tone sent from the IP side. The fax terminal that detects the fax tone automatically switches to the transparent transmission mode without the SIP negotiation.

One problem currently exists in the auto-switching fax flow: If the DSP channel originally works in the G.729 mode for the voice service, and is now switched to the G.711 transparent transmission mode when the fax tone is detected, the G.711 voice packet may not be recognized. This is because the DSP channel of the calling party stills works in the G.729 mode. Therefore, the DSP chip is required to be able to receive G.711 packets when working in the G.729 or other coding modes. The prerequisite remains that the DSP chip should detect and report the fax tone sent from the IP side.

Flow of the Auto-Switching T.38 Fax

The working principle of this fax flow is the same as the working principle of the auto-switching transparent transmission fax. The difference is that, after the fax tone is detected, the DSP channel is enabled in the T.38 mode instead of the transparent transmission mode.

SIP-Based MoIP

In terms of service flow, the modem service is similar to the transparent transmission fax service, and can also be classified as auto-switching and negotiated-switching.

The modem service in the negotiated-switching transparent transmission mode can be presented in three ways.

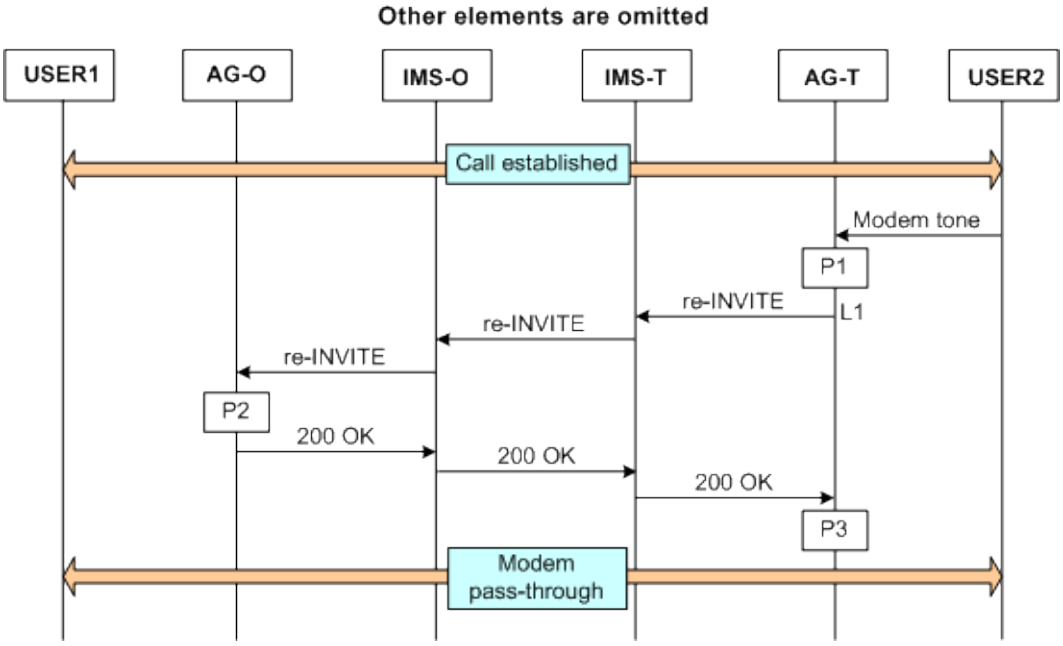
- Presented as a=modem. This is a G.711 transparent transmission modem mode proposed by China Telecom.
- Presented as a=silenceSupp:off. This is a G.711 transparent transmission modem mode defined in the *draft-IETF-sipping-realtimefax-01.txt*.
- Presented as a=gpmd:99 vbd=yes. This is a VBD mode defined in the ITU-T V.152.

The method actually applied depends on the parameters configured.

Flow of the Negotiated-Switching Modem Service

Figure 23-39 shows the flow of the negotiated-switching modem service.

Figure 23-39 Flow of the negotiated-switching modem service



- P1: AG-T first detects the modem tone, and then sends the re-INVITE message to the AG (AG-O) to which the calling party is connected.
- L1: The SDP message contained in the re-INVITE message has three types, corresponding to the three preceding presentations of the negotiated-switching transparent transmission mode. The specific transparent transmission modem mode must be configured on the AGs.
- P2: AG-O receives the re-INVITE message. Then, AG-O generates the 200 message and sends the message to AG-T.
- P3: AG-T receives the 200 OK message, and also enables the DSP channel in the fax or modem mode.

Auto-Switching Modem Mode

In this mode, after the AG detects the modem tone, the AG automatically switches the DSP channel to the VBD mode without notifying the IMS or the peer device.

Generally, the called modem detects the modem tone on the TDM side first, and the calling modem detects the modem tone sent from the IP side. The modem that detects the modem tone automatically switches to the VBD mode without the SIP negotiation.

SIP Trunking

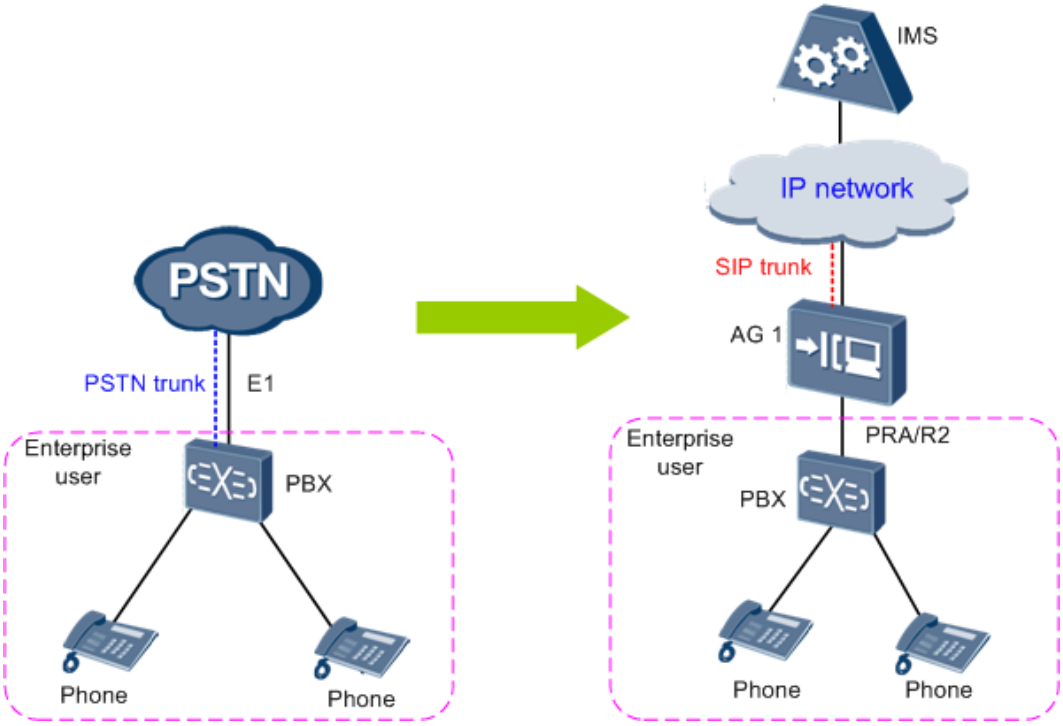
This section describes the SIP trunking feature and its typical usage scenarios.

Background

The traditional private branch exchanges (PBXs) of government, enterprise, and business users connect to carrier networks using PSTN trunk lines (E1 lines). The All-IP network upgrade and reconstruction urgently demands an IP trunk technology to replace the traditional PSTN trunk technology. Then SIP trunking technology can meet such requirements,

connecting government, enterprise, and business PBXs to carrier networks for voice and multimedia services. Figure 23-40 shows the network reconstruction for SIP trunking.

Figure 23-40 Network reconstruction for SIP trunking



Basic Concepts

- Trunk: a physical communication line connecting two switching systems for carrying media stream signals, such as voice, data, and video signals.
- The SIP trunking technology uses SIP to connect government, enterprise, and business PBXs to carrier networks over an IP network.
- SIP trunking allows a PBX to register with the IMS in one of the following modes:
  - Separate number registration: In this mode, the PBX uses a separate number but not a wildcard number to register with the IMS.
  - Wildcard number registration: When the PBX successfully registers with the IMS using a wildcard number, such as 2878*, all numbers with prefix "2878" are successfully registered on the IMS.
  - Agent group registration: In this mode, multiple numbers are added to one group and the PBX uses one of the numbers in this group or a separate group number to register with the IMS. If the registration is successful, all numbers in this group are successfully registered on the IMS. This registration mode reduces registration message exchanging between the IMS and the AG connected to this PBX.
  - No registration: In this mode, the PBX supports call originating without registration. This mode is used if the access network is reliable.
- SIP trunking supports call originating in one of the following modes:
  - Direct dialing in (DDI) calling: In this mode, the PBX requires multiple extension numbers and uses a separate number to register with the IMS; or the PBX uses a wildcard number to register with the IMS. The extension phones connected to the

PBX can call each other using either long or short numbers. External users can call an extension phone by dialing its long number. User experience in DDI calling mode is the same as that in POTS access mode.

- Global dialing number (GDN) calling: In this mode, the PBX requires only a switchboard number, also called "pilot number". The extension phones connected to the PBX call each other using short numbers. When an external user wants to call an extension phone, the user must dial the switchboard number up and the switchboard transfers to the call to the extension phone. In GDN calling mode, the number displayed for the external user is not the number of the extension phone where the call is initiated but the number of the switchboard number. The external user cannot call this extension phone in callback mode.
- Line hunting: In this mode, multiple E1 lines connected to the PBX are added to one hunting group and one number is configured for this group. One hunting group can have one or multiple group numbers. When the PBX registers with the IMS, the AG adds these group numbers to one agent group for registration. For details about line hunting call process, see Line Hunting.
- Call routing identified by **tgrp** and **trunk-context**: This mode is used if call routing and charging cannot be implemented using only phone numbers. **tgrp** and **trunk-context** are defined in RFC 4904. They are only used in the Request URI and Contact URI, functioning as **useinfo** in the SIP URI or **par** in the TEL URI. The two parameters must be used in couple. Otherwise, AG considers that the URI does not carry this group of parameters. Huawei AG devices use these two parameters in hunting groups for call routing and charging.

Typical Usage Scenarios

Scenario 1: DDI calling

Usage scenario: As shown in Figure 23-41, the enterprise user uses one PBX to connect to 30 extension phones and the PBX connects to the AG using one PRA port. The 30 extension phones use different phone numbers, 28780001 through 28780030.

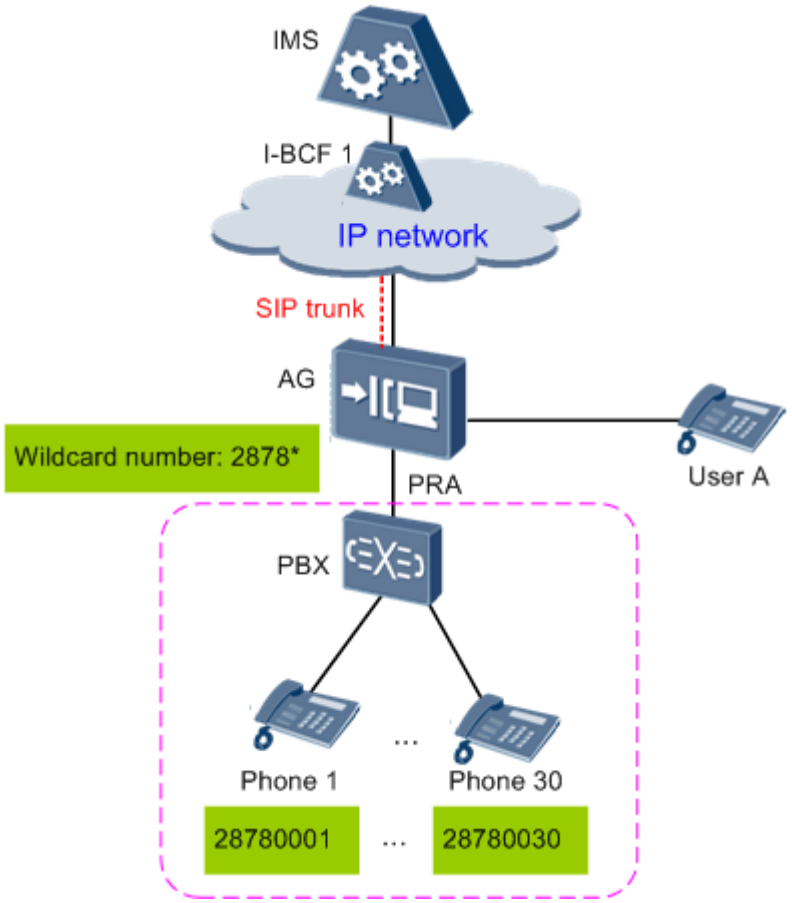
Configuration: Wildcard number 2878* is configured for the PRA port and the PBX uses this wildcard number to register with the IMS.

NOTE
If the IMS does not support the registration using a wildcard number, configure 1 primary number and 29 extension numbers for the PBX. All the 30 numbers require a separate number registration.

Call description:

- When user A dials number 28780001 up, phone 1 is called without requiring call transferring from the switchboard.
- When phone 1 calls user A, the number displayed for user A is 28780001, the number of phone 1.
- Phone 1 can call phone 30 using either long number 28780030 or short number 0030 of phone 30.

Figure 23-41 DDI calling



Scenario 2: GDN calling

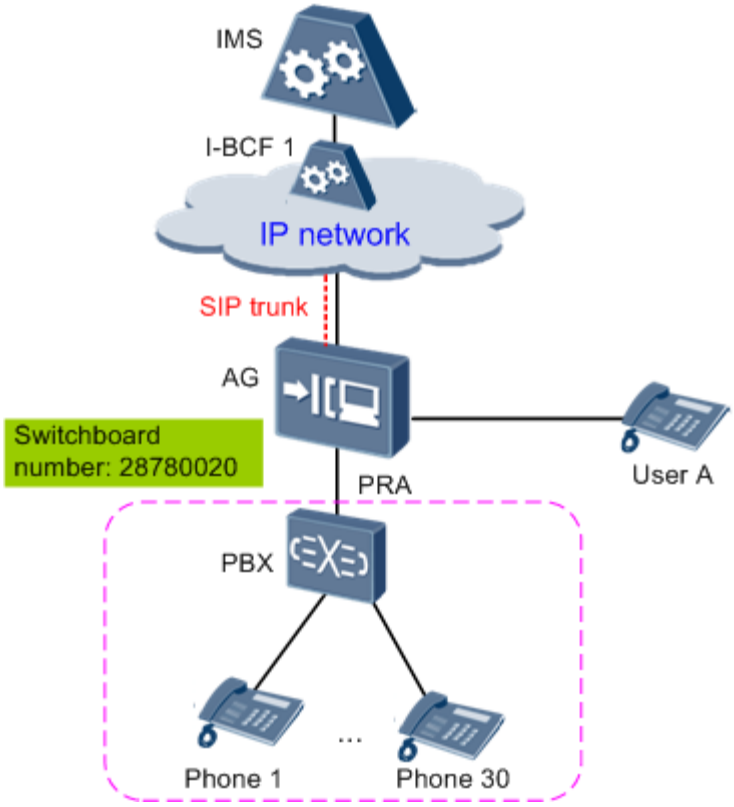
Usage scenario: As shown in Figure 23-42, the enterprise user uses one PBX to connect to 30 extension phones and the PBX connects to the AG using one PRA port. The external number of the 30 extension phones is 28780020.

Configuration: Number 28780020 is configured for the PRA port and the PBX uses this number to register with the IMS in separate number registration mode.

Call description:

- When user A calls number 28780020, the AG routes the call to the PBX. Then, the PBX uses the interactive voice response (IVR) function to automatically transfer the call to the desired extension phone, or uses the switchboard to manually transfer the call to the desired extension phone.
- When an extension phone calls user A, the number displayed for user A is 28780020.
- Phone 1 can call phone 30 only using short number 0030 of phone 30.

Figure 23-42 GDN calling



Scenario 3: line hunting calling

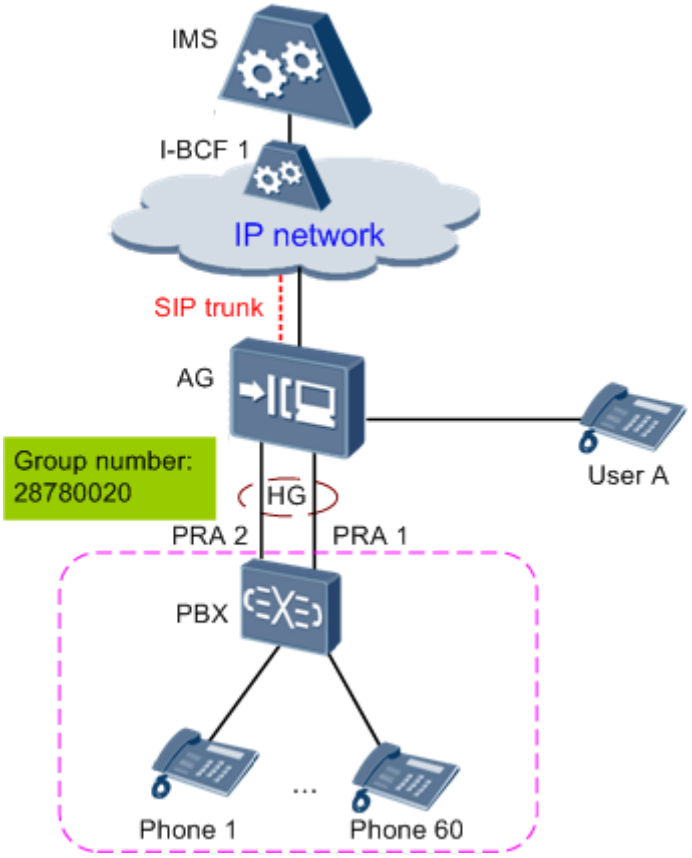
Usage scenario: As shown in Figure 23-43, the enterprise user uses one PBX to connect to 60 extension phones and the PBX connects to the AG using two PRA ports: PRA 1 and PRA 2. The external number of the 60 extension phones is 28780020.

Configuration: PRA 1 and PRA 2 are added to hunting group **HG**, number 28780020 is configured for this hunting group, and the PBX uses this number to register with the IMS.

Call description:

- When user A calls 28780020, the AG routes the call to hunting group **HG** and this group selects an idle PRA port based on hunting policies for call incoming. Then, the PBX uses the IVR function to automatically transfer the call to the desired extension phone, or uses the switchboard to manually transfer the call to the desired extension phone.
- When phone 1 calls user A, the PBX selects an idle PRA port for call outgoing. The number displayed for user A is 28780020.
- Phones 1 through 60 can call each other using only short numbers.

Figure 23-43 Line hunting calling

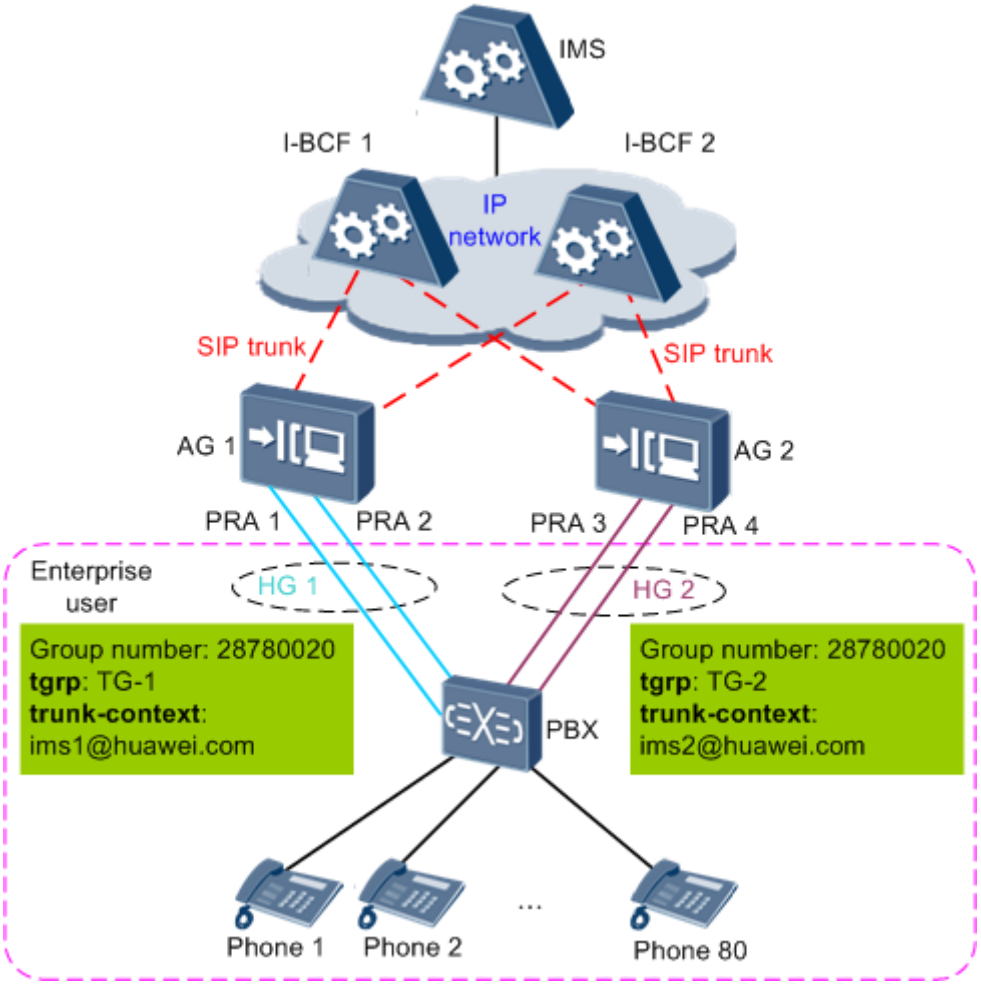


Scenario 4: PBX dual-homing calling (call routing identified by tgrp and trunk-context)

- Usage scenario: As shown in Figure 23-44, the enterprise user uses one PBX to connect to 80 extension phones. The external number of the 80 extension phones is 28780020. Because of high reliability requirements, the PBX uses two PRA ports to connect to two AGs, respectively, for redundancy backup on the access side; both AGs connect to I-BCF 1 and I-BCF 2, respectively, for dual homing on the core network side. This networking improves call reliability. However, the two AGs use the same number. Therefore, calls cannot be routed based on only the phone number. To resolve this issue, **tgrp** and **trunk-context** are added for differentiating between hunting groups for different AGs.
- Configuration:
  - For AG 1:
    - PRA 1 and PRA 2 on AG 1 are added to hunting group **HG1**.
    - Number 28780020 is configured for this hunting group.
    - The values of **tgrp** and **trunk-context** are **TG-1** and **ims1@huawei.com**, respectively.
  - For AG 2:
    - PRA 3 and PRA 4 on AG 2 are added to hunting group **HG2**.
    - Number 28780020 is configured for this hunting group.
    - The values of **tgrp** and **trunk-context** are **TG-2** and **ims2@huawei.com**, respectively.
- Call description:

- When an external user calls number 28780020, the IMS routes the call to the desired AG based on the phone number, **tgrp**, and **trunk-context**. The AG selects an idle PRA port based on hunting policies for call incoming. Then, the PBX uses the IVR function to automatically transfer the call to the desired extension phone, or uses the switchboard to manually transfer the call to the desired extension phone.
- When an extension phone calls an external user, the PBX selects an idle PRA port for call outgoing. The phone number of the calling party is mapped to contact header field URI and carries the **tgrp** and **trunk-context** parameters for the hunting group. Then, the IMS charges the call based on the **tgrp** and **trunk-context** parameters.

Figure 23-44 PBX dual-homing calling



Standards and Protocols Compliance

- RFC 4904: representing trunk groups in tel/sip uniform resource identifiers (URIs)
- RFC 3261: Session Initiation Protocol
- RFC 3966: tel URI for phone numbers
- ETSI TS 182 025: Telecommunications and Internet Converged Services and Protocols for Advanced Networking (TISPAN); business trunking; architecture and functional description

Line Hunting

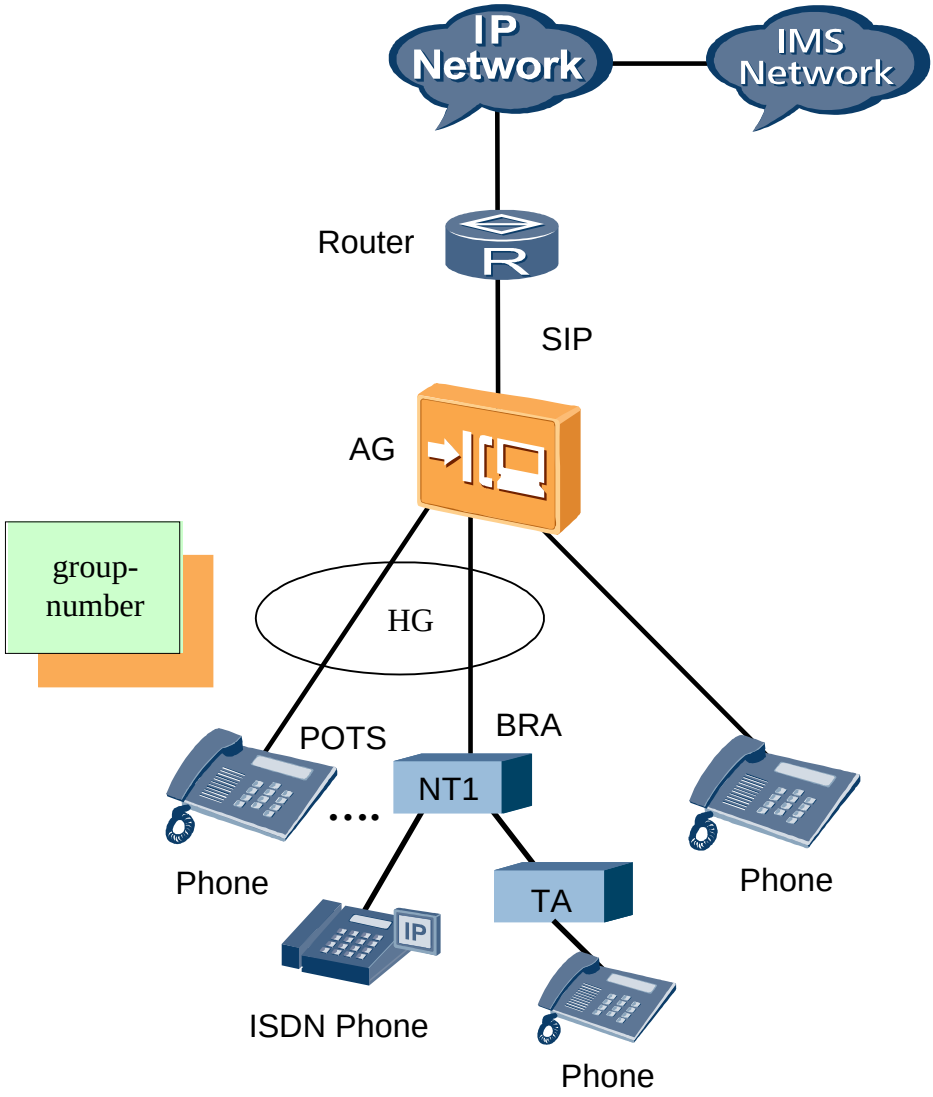
Line hunting is a feature that allows a group of ports to share a group of called party numbers by specifying a hunting group and hunting policy. Only the Session Initiation Protocol (SIP) supports this feature.

Basic Concepts

Figure 23-45 shows a hunting group. The following provides the basic concepts about line hunting.

- Hunting group: a group composed of multiple members that share a group of called party numbers, for example, the HG shown in Figure 23-45
- Group member: a port in a hunting group. A group member can be a POTS port, a BRA port, a PRA port or a child hunting group. A child hunting group is a sub group of a parent hunting group. A child hunting group has the same structure as its parent hunting group. After a child hunting group is added into a parent hunting group, all the members of the child hunting group become the members of the parent hunting group.
- Group number: a called party number shared by members of a hunting group. A group number can be a group numbering reduced (GNR) or a direct dial number (DDN). When the access gateway receives an incoming call and if the incoming call number is a group number of a hunting group, the AG hunts for a group member according to the hunting policy. A hunting group has one or more group numbers, but a group number belongs to only one hunting group.
  - GNR: a prefix of a wildcard number. For example, if the wildcard number is 024545*, the GNR is 024545. For simplicity, "024545*" is called a GNR. If the group number of a hunting group is a GNR, the AG determines that an incoming call number is the group number of the hunting group as long as the first several digits of the incoming call number are identical to the prefix of the GNR.
  - DDN: a specific number instead of a wildcard number
- Hunting policy: a policy used to select a member of a hunting group as the called party for an incoming call. Hunting policies include sequential hunting, circular hunting, and circular hunting by weight. For details, see Hunting Policies.
- Alternative line hunting: defines the call release mode for an ISDN port in a hunting group. Only ISDN ports support alternative line hunting. For details, see Alternative Line Hunting.

Figure 23-45 Hunting group



Hunting Policies

Available hunting policies include sequential hunting, circular hunting, and circular hunting by weight.

Item	Sequential Hunting	Circular Hunting	Circular Hunting by Weight
order	The **order** value determines the priority of a member port. When the AG receives an incoming call, the member port with the **order** value 1 takes precedence. If the member port with the **order** value 1 is busy	The **order** value determines the neighbor relationship between member ports. For example, the ports with the order values 1 and 2 are considered as neighbor ports. The neighbor port next to	The meaning is the same as the **order** in circular hunting.

Item	Sequential Hunting	Circular Hunting	Circular Hunting by Weight
	or faulty, the member port with the **order** value 2 is selected. If the member port with the order value 1 is busy or faulty, the member port with the order value 2 is selected, and so on.	the previously selected port is the first choice when the AG is hunting for a group member. Specifically, when the AG receives the first incoming call, the member port with the **order** value 1 takes precedence. If the port is busy or faulty, the member port with the **order** value 2 is selected. If the member port with the **order** value 2 is successfully selected as the called party, the member port with the **order** value 3 is selected when the AG receives the second incoming call.	
weight	–	–	The **weight** value determines the number of times that a member port is selected as the called party. When the **weight** value of each member port in a hunting group is the same, the circular hunting by weight functions the same as the circular hunting. In circular hunting by weight, the **weight** value decreases by 1 each time after a member port is successfully selected as the called party. If there are still member ports whose **weight** values is not 0 after all member ports in a hunting group are selected once as the called parties, the AG preferentially selects the member ports with non-0 **weight** values as the called parties according to the circular hunting policy.

NOTE
Members of a hunting group can have the same **order** value.

- When a child group and a port have the same **order** value, the port takes precedence.
- When two or more ports have the same **order** value, the ports are selected according their subrack IDs/slot IDs/ port IDs. Specifically, the port in the subrack with the smallest ID takes precedence. If these ports are in the same subrack, the port on the board with the smallest slot ID takes precedence. If these ports are on the same board, the port with the smallest ID takes precedence.
- When two or more sub groups have the same **order** value, they are selected according to the numerical and alphabetical sequences of their names.

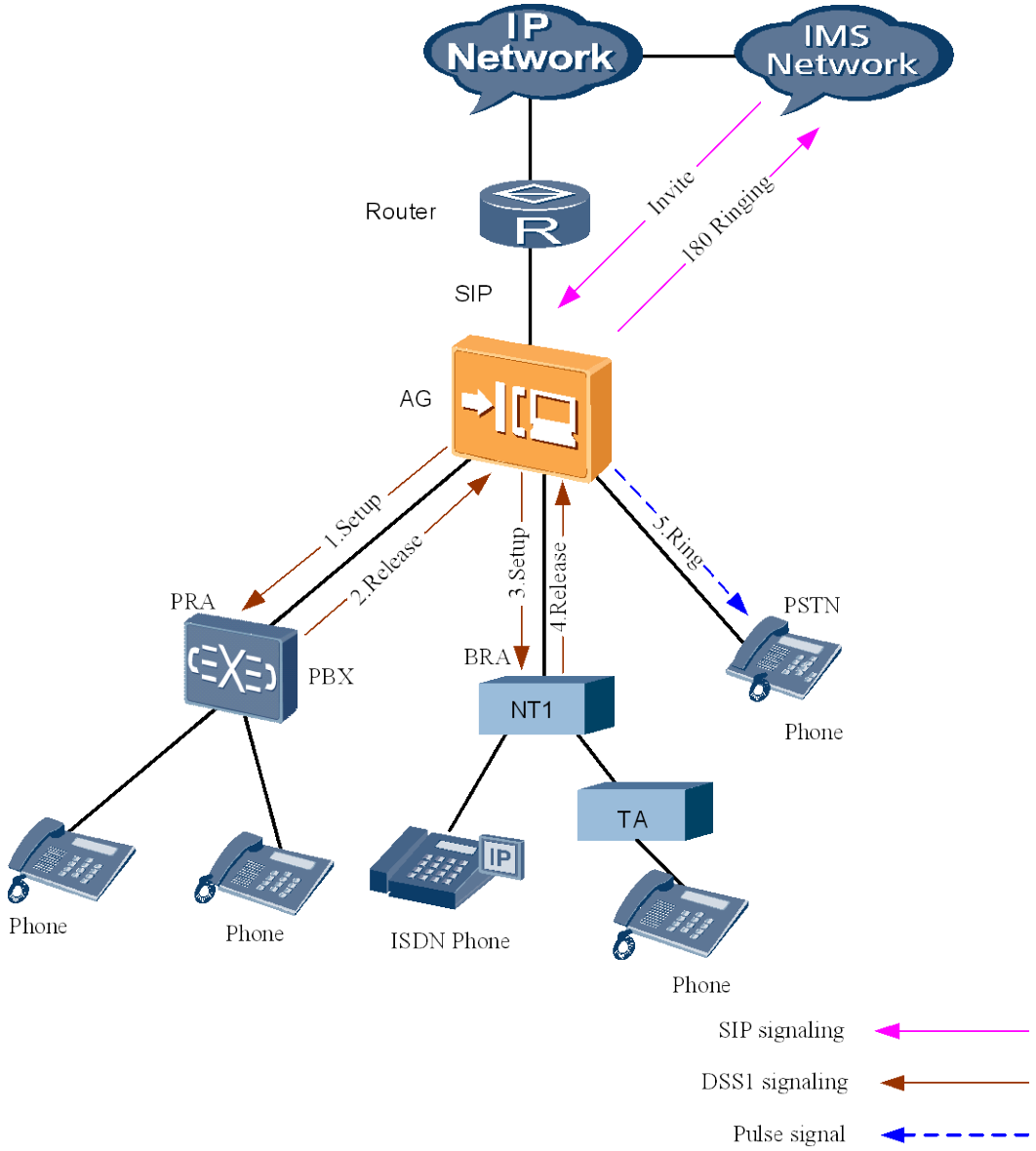
Alternative Line Hunting

Assume that the AG selects an ISDN port as the called party for an incoming call and issues a call request. If the AG receives a call failure message before receiving any response and the Q.850 cause code carried in the call failure message (specifically, Release message) cannot be found in the Q.850 cause code table configured for the hunting group, the AG does not release the call but continues to hunt for another available group member and issues the call request. On the contrary, if the Q.850 cause code is found in the Q.850 cause code table, or if the AG receives a response before receiving any call failure message, the AG releases the call. This procedure is called alternative line hunting. For the definition of the Q.850 cause code, see *ITU-T Q.850*.

Assume that:

- The PRA port, BRA port, and POTS port (as shown in Figure 23-46) are members of the same hunting group.
- The hunting policy of the hunting group is sequential hunting. The **order** value of the PRA port is set to 1. The **order** value of the BRA port is set to 2. The **order** value of the POTS port is set to 3.
- The PRA, BRA, and POTS ports are in the idle state.
- The Q.850 cause code carried in the Release message cannot be found in the Q.850 cause code table configured for the hunting group.

Figure 23-46 Alternative line hunting



The called party number of an incoming call is a group number. When receiving an Invite message sent by the IP multimedia subsystem (IMS), the AG hunts for a member among multiple group members. The alternative line hunting procedure is as follows:

1. The AG issues a Setup message to the PRA port.
2. The PBX directly replies to the AG with a Release message. The AG cannot find the Q.850 cause code (carried in the Release message) in the Q.850 cause code table configured for the hunting group, and then continues to hunt for a member.
3. The AG issues a Setup message to the BRA port.
4. The BRA replies to the AG with a Release message. The AG cannot find the Q.850 cause code (carried in the Release message) in the Q.850 cause code table configured for the hunting group, and then continues to hunt for a member.

5. The AG issues ringing signals to the POTS port, and replies to the IMS with a 180 Ringing signaling message. Then the POTS port is selected as the called party.

23.5.4 SIP Value-added Services

SIP value-added services provide more services with easier operations for users and help carriers provide various and flexible services for users. These services improve carriers' competitiveness and user satisfaction.

List of Value-added SIP Services

Table 23-13 describes SIP value-added services.

Table 23-13 SIP value-added services

Service	Definition	POTS User	ISDN User
Call hold (CH)	This service allows a user to temporarily place an established call on hold. Media streams between the calling and called parties are not sent any more; however, the call resource is not released. This call can be resumed if the user requires. This function works when a user needs to make a new outgoing call or answer a new incoming call in an ongoing call.	Y	Y
Call waiting (CW)	If this callee-side service is provisioned and activated and the user who is engaged in an ongoing call is notified of a new incoming call, this new incoming call will be in waiting state and the waiting party hears the CW tone. The user can either accept, reject, or ignore this new incoming call.	Y	Y
Explicit communication transfer (ECT)	This service allows user A who is communicating with user B to transfer the call to user C so that a call is set up between user B and user C.	Y	Y
Hotline	This service functions in this way: If the user who registers this service does not dial a number in a specified period (such as 5s) after picking up the phone, the system automatically connects this user to a fixed number (hotline number). The hotline service can be classified into ordinary hotline and immediate hotline according to the no-dial interval. • Ordinary hotline: Also called delay hotline. If the user does not dial a number in a specified period after picking up the phone, the system automatically connects this user to the hotline number. • Immediate hotline: After the user picks up the phone, the system automatically connects this	Y	N

Service	Definition	POTS User	ISDN User
	user to the hotline number.		
Call hold with three parties	With this service, a user (namely, user A) places the communicating user (namely, user B) on hold and initiates a new call (namely, user C). Then, user A can communicate with user B and user C alternatively or terminate either of the calls.	Y	Y
Three-party conference service (3PTY)	This service provides one call connection for multiple users, that is, it allows 3 users to communicate with each other in the same call.	Y	Y
Conferencing	The conferencing party can apply for conference resources, and add or delete conference participants. Conference participants can leave conferences. Compared with the 3PTY service, the IP multimedia subsystem (IMS) controls multimedia resource function controller (MRFC) to implement N-party conferencing, allowing more than 3 parties, and up to 64 parties, to join a conference. The way of setting up a multi-party conference is the same as that of setting up a 3PTY.	Y	Y
Message waiting indication (MWI)	This is a message prompt service, with which, the voice mail system (VMS) or unified message system (UMS) notifies users of changed number or status of messages, including emails, short messages, faxes, and leaving messages.	Y	Y
Anonymous call service	This service does not allow the number of the calling party who registers this service to be presented to the called party.	Y	Y
Malicious call identification (MCID)	A user who registers this callee-side service can identify the calling number if the user receives a malicious call or unsolicited call, such as a prank call or a phishing call.	Y	Y
Emergency call	A user can make an emergency call if the connected port is in remote block state and the SIP proxy server is normal and the dialed number matches the emergency digitmap.	Y	N
Call forwarding (CF)	The system supports the following 3 call forwarding services: • Call forwarding unconditional (CFU): A callee-side service, with which, a user can unconditionally forward all incoming calls to a designated forwarded-to number. • Call forwarding busy (CFB): A callee-side	Y	Y

Service	Definition	POTS User	ISDN User
	service, with which, a user can forward all incoming calls to a designated forwarded-to number when the subscriber is busy on another call. • Call forwarding no reply (CFNR): A callee-side service, with which, a user can forward all incoming calls to a designated forwarded-to number if the calls are not answered within a preset period.		
Distinctive ringing	A callee-side service, with which, a user can set different ringing tones for different calling parties.	Y	N
Call release control	Call release control is classified into calling party release, called party release, and first party release. They are the 3 modes for call release. • Calling party release: A call is not released if the called party hangs up the phone but the calling party does not. In this case, if the called party picks up the phone again before the timer for calling party release times out, the call is connected and the two parities communicate with each other further. A call is released if the calling party hangs up the phone. • Called party release: A call is not released if the calling party hangs up the phone but the called party does not. In this case, if the calling party picks up the phone again before the timer for called party release times out, the call is connected and the two parities communicate with each other further. A call is released if the called party hangs up the phone. • First party release: A call is released if either of the calling party and called party hands up the phone.	Y	N
Multi-number service	Extended phone numbers are supported for SIP user configurations. The basic number is configured when a user is added or modified. A SIP user supports one basic phone number and multiple extended phone numbers. A SIP user can configure only extended phone numbers (does not configure the basic number). An extended phone number can be a local phone number or global phone number. User registration and subscription can be initiated and managed by phone number. When a SIP user makes a call as caller, the first number functions as the active number. If registration fails, the first number in the following numbers is used for registration again. The rest may be deduced by	Y	N

Service	Definition	POTS User	ISDN User
	analogy and the phone number for the first successful registration is bound to the call. When a SIP user makes a call as callee, the SIP access gateway (AG) selects a phone number for the call according to the URI carried in the INVITE message.		
Calling line identification presentation (CLIP)	A callee-side service allows the number of the calling party to be presented to the called party.	Y	Y
Calling line identification restriction (CLIR)	A caller-side service prevents the number and name of the calling party from being presented to the called party who does not register RIO. For RIO, see the following service.	Y	Y
Calling line identification restriction override (RIO)	A callee-side service allows the number of the calling party to be presented to the called party even if the calling party registers CLIR.	Y	Y

Call Hold

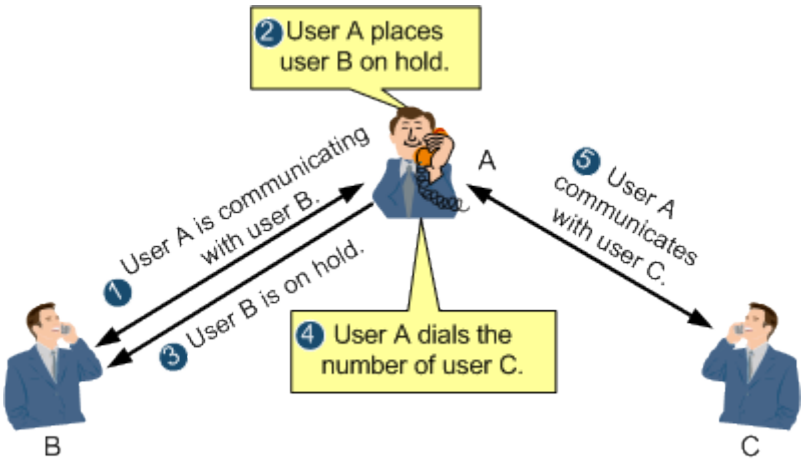
This topic describes the definition and principle of the call hold (CH) service.

Definition

This service allows a user to temporarily place an established call on hold. Media streams between the calling and called parties are not sent any more; however, the call resource is not released. This call can be resumed if the user requires. This function works when a user needs to make a new outgoing call or answer a new incoming call in an ongoing call.

Assuming that user A registers the CH service, user B is placed on hold, and user C is the third party, the principle of the CH service is as shown in Figure 23-47.

Figure 23-47 Principle of the CH service



1. User A is communicating with user B.
2. User A wants to communicate with user C and therefore user A presses the hookflash button on the terminal to place user B on hold.
3. User B is on hold and hears the CH tone.
4. User A dials the number of user C.
5. User C answers the call and communicates with user A.

NOTE

The preceding figure shows one application scenario. In another scenario, when user C dials user A, user A can place user B on hold and answer the call from user C.

Benefit

Beneficiary	Benefits
Carrier	The CH service supplements the value-added services of carriers and helps improve the ratio of successful call connections.
User	A user can place the ongoing call on hold and then resumes this call if required, reducing the number of dials and facilitating call making.

Standard Compliance

ETSI TS 183 010

3GPP TS 24610

Call Waiting

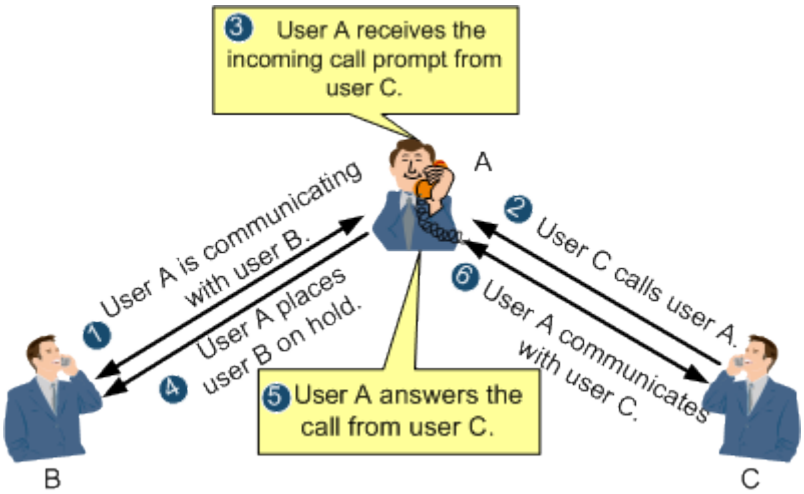
This topic describes the definition and principle of the call waiting (CW) service.

Definition

If this callee-side service is provisioned and activated and the user who is engaged in an ongoing call is notified of a new incoming call, this new incoming call will be in waiting state and the waiting party hears the CW tone. The user can either accept, reject, or ignore this new incoming call.

Assuming that user A registers the CW service, user B is placed on hold, and user C is the third party (CW party), the principle of the CW service is as shown in Figure 23-48.

Figure 23-48 Principle of the CW service



1. User A is communicating with user B.
2. User C calls user A.
3. User A receives the incoming call prompt from user C.
4. User A wants to communicate with user C and places user B on hold.
5. User A answers the call from user C.
6. User A communicates with user C.

Benefit

Beneficiary	Benefits
Carrier	The CW service helps improve the ratio of successful call connections, enhances carriers' competitiveness and provides carriers with new revenue growth potential.
User	• The CW service prevents a caller from dialing a number for multiple times and improves the ratio of successful call connections. • The CW service reduces the number of missed calls and the possibility of missing expected calls. The CW service can be cancelled temporarily so that the ongoing call is not affected.

Standard Compliance

ETSI TS 183 036

ETSI 300 102-1

Explicit Communication Transfer

This topic describes the definition and principle of the explicit communication transfer (ECT) service.

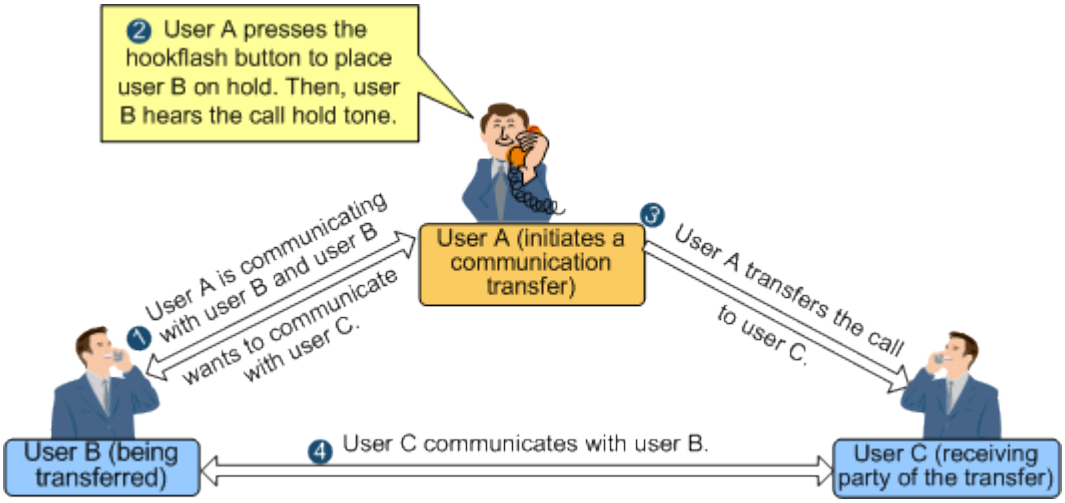
Definition

This service allows user A who is communicating with user B to transfer the call to user C so that a call is set up between user B and user C. This service is classified into two types: explicit transfer and blind transfer.

- Explicit transfer: When user C is in ringing state or is communicating with user A, user A initiates ECT and then transfers user B to user C. In explicit transfer, the call to user A is released only after user B is successfully transferred to user C.
- Bland transfer: When user A initiates ECT, the system directly releases the call to user A. In blind transfer, user A does not know whether user B can communicate with user C.

Assuming that user A registers the ECT service and initiates a communication transfer, user B is transferred, and user C is the receiving party of the transfer, the principle of the ECT service is as shown in Figure 23-49.

Figure 23-49 Principle of the ECT service



1. User A is communicating with user B and user B wants to communicate with user C.
2. User A presses the hookflash button to place user B on hold. Then, user B hears the call hold tone and user A hears the special dial tone.
3. User A dials the number of user C.
4. The phone of user C is ringing. User A initiates a communication transfer and is released, and user B hears the ring back tone.
5. User C answers the call and communicates with user B.

NOTE

If the communication transfer fails, the system re-calls user B to make user A communicate with user B.

Benefit

Beneficiary	Benefits
Carrier	The ECT service enhances carriers' competitiveness and provides carriers with new revenue growth potential.
User	The ECT service allows a user to easily transfer a call to another user so that the desired user can answer the call.

Standard Compliance

ETSI TS 183 010

3GPP TS 24610

Three-party Conference Service

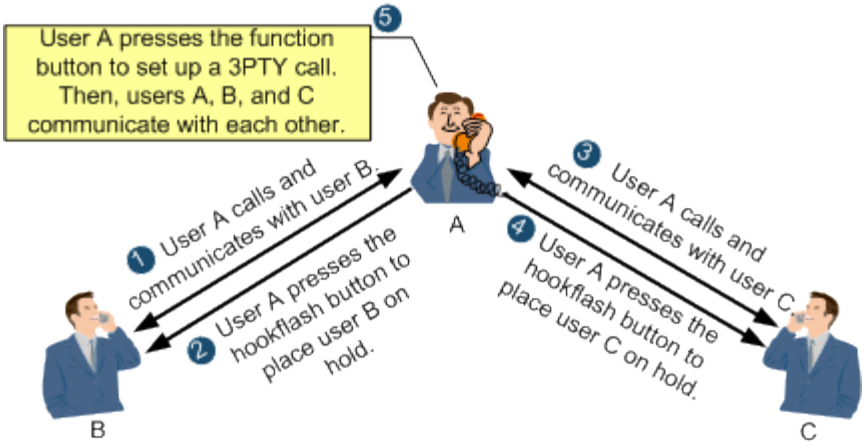
This topic describes the definition and principle of the three-party conference service (3PTY) service.

Definition

This service provides one call connection for multiple users, that is, it allows 3 users to communicate with each other in the same call.

Assuming that user A registers the 3PTY service, users B and C are ordinary users, and user A sets up a call with users B and C, the principle of the 3PTY service is as shown in Figure 23-50.

Figure 23-50 Principle of the 3PTY service



1. User A calls and communicates with user B.

2. User A presses the hookflash button to place user B on hold. Then, user B hears the call hold tone.
3. User A calls and communicates with user C.
4. Users A presses the hookflash button to place user C on hold. Then, user C hears the call hold tone.
5. User A presses the function button to set up a 3PTY call. Then, users A, B, and C communicate with each other.

NOTE
Generally, the function button is **3**.

Benefit

Beneficiary	Benefits
Carrier	The 3PTY service supplements the value-added services of carriers, provides multi-party calls for users, and improves user satisfaction.
User	The 3PTY service enables a user to initiate a 3PTY call for discussion, facilitating user communication and improving communication efficiency.

Standard Compliance

- ETSI TS 183043
- ETSI TS 183005
- 3GPP TS 24147
- 3GPP TS 24605
- 3GPP TS 23218-630
- 3GPP TS 24610

23.5.5 SIP Reference Standards and Protocols

- RFC 3262: Reliability Of Provisional Responses in the Session Initiation Protocol (SIP)
- RFC 3263: SIP Locating SIP Servers

23.6 MGCP Voice Feature

This topic describes the MGCP protocol and the working principle of MGCP application in VoIP, MoIP and FoIP.

23.6.1 Introduction to the MGCP Feature

Definition

Defined by IETF, MGCP is a protocol that specifies a call control mechanism in which call control and service bearing are separated. Call control is independent of the media gateway

(MG) and is processed by the MGC. Therefore, MGCP is actually a master-slave protocol. The MG establishes various service connections under the control of the MGC.

MGCP provides the following commands:

1. NotificationRequest: The MGC sends this command to request the MG to detect a specified event, such as an offhook event or onhook event. After detecting such an event, the MG notifies the MGC. Through this command, the MGC can also instruct the MG to play signal tones, such as the dial tone and busy tone.
2. Notify: After the MG detects the specified event as instructed by the MGC, the MG sends this command to notify the MGC of the detected event.
3. CreateConnection: The MGC sends this command to instruct the MG to create a media connection. The command contains the instruction or suggestion on the bearing parameters and connection parameters.
4. ModifyConnection: The MGC sends this command to instruct the MG to modify the bearing parameters and connection parameters of an established media connection.
5. DeleteConnection: The MGC sends this command to instruct the MG to delete an established media connection. The MG can also voluntarily delete a connection. This means that, when the MG discovers that system resources are insufficient or the system is faulty, the MG can delete the connection and at the same time send this command to notify the MGC. Therefore, this command is bi-directionally available between the MGC and the MG.
6. AuditEndpoint and AuditConnection: The MGC sends the commands to check the status of a specified endpoint and connection.
7. RestartInProgress: The MG sends this command to notify the MGC that the MG or a certain endpoint managed by the MG is not available or is becoming available. This command is usually triggered by a system fault or restart.

MGCP also provides the following features:

1. Encoding in the text format
2. Adopting the Session Description Protocol (SDP) to describe the connection parameters of the media stream
3. Introducing the concept of event package
4. Adopting the wildcard to describe endpoints and events

Purpose

MGCP solves the internal problems of MG and media devices, thus realizing an open distributed system which is formed by the MG and media devices.

In the MGCP mechanism, the MG and media devices are separated into two logically independent parties, the MG and the MGC, which communicate through MGCP. The MG processes the user plane, and the MGC processes the control plane and controls the actions of the MG. In other words, the MG acts under the control of the MGC.

23.6.2 MGCP Principles

MGCP-Based VoIP

Figure 23-51 illustrates the principle of the call establishment and release in the MGCP-based VoIP service.

Figure 23-51 Principles of the call establishment and release in the MGCP-based VoIP service

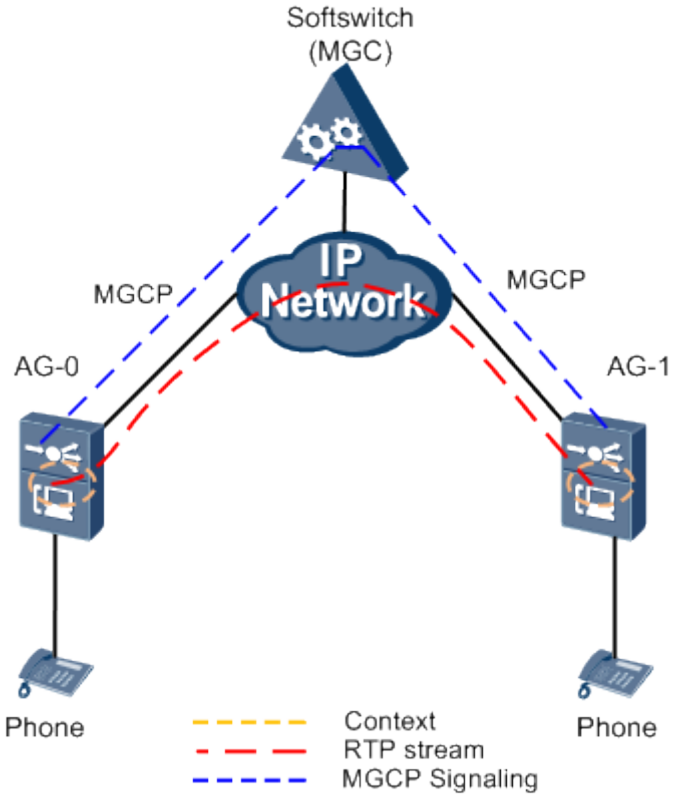
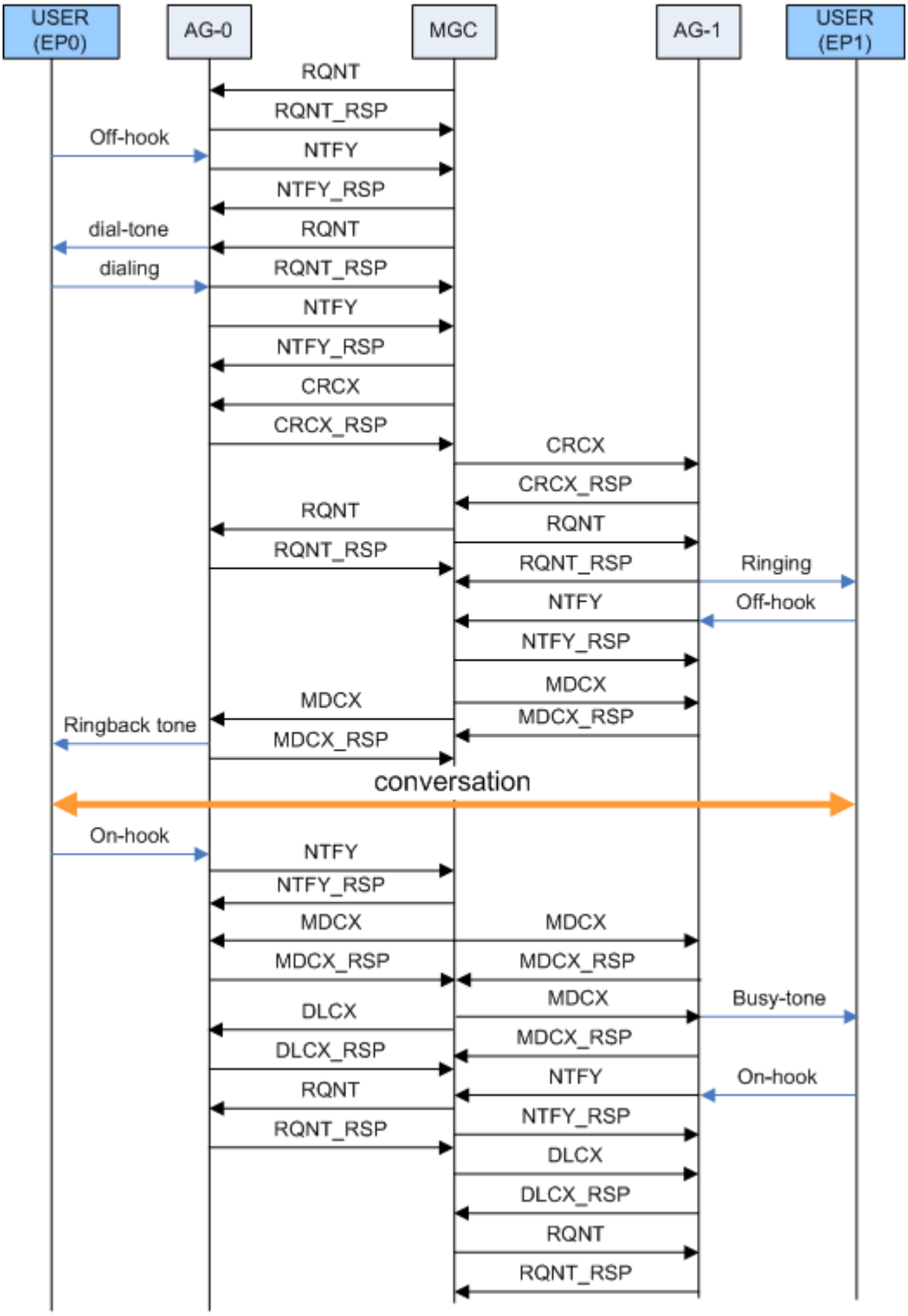


Figure 23-52 illustrates the basic flow of a call establishment and release process.

Figure 23-52 MGCP-based call flow



1. AG-0 detects the offhook of EP0, and notifies the MGC of the offhook event through the Notify command.

2. After the MGC receives the offhook event, the MGC sends a digitmap to AG-0, requests AG-0 to play the dial tone to EP0, and at the same time checks for the digit collection event.
3. User EP0 dials a telephone number, and AG-0 collects the digits according to the digitmap issued by the MGC. Then, AG-0 reports the result of digit collection to the MGC.
4. The MGC sends the CRCX (CreateConnection) command to AG-0 requesting AG-0 to create a connection at endpoint EP0.
5. AG-0 allocates resources for creating this connection and sends a response to the MGC. The response contains the session description that provides the necessary information for the peer end to send the packet to AG-0, such as the IP address and UDP port number.
6. The MGC sends the CRCX command to AG-1 requesting AG-1 to create a connection at endpoint EP1.
7. AG-1 allocates resources for creating this connection and sends a response to the MGC. The response contains the session description that provides the necessary information for the peer end to send the packet to AG-1, such as the IP address and UDP port number.
8. AG-1 detects the offhook of EP1, and sends the Notify command to the MGC. The softswitch (MGC) sends the MDCX (ModifyConnection) command to stop the ring back tone of EP0 and the ringing of EP1.
9. The MGC sends the session description of AG-1 to EP0 through the MDCX command. Then, the conversation is set up between EP0 and EP1.
10. AG-0 detects the onhook of EP0, and notifies the MGC of the onhook event through the Notify command.
11. The MGC sends the MDCX command to AG-0 and AG-1 respectively to modify the RTP resource to receive-only.
12. The MGC sends the MDCX command to AG-1 requesting AG-1 to play the busy tone to EP1, and at the same time checks for the onhook event.
13. The MGC sends the DCLX (DeleteConnection) command to AG-0, requesting AG-0 to release the resources that are occupied by the call of EP0.
14. AG-1 detects the onhook of EP1, and notifies the MGC of the onhook event through the Notify command.
15. The MGC sends the DCLX command to AG-1, requesting AG-1 to release the resources that are occupied by the call of EP1.
16. The call between EP0 and EP1 is terminated, and all the resources occupied by the call are released.

MGCP-Based MoIP

MoIP refers to the modem service provided on the IP network or between the IP network and the traditional PSTN network. According to different control devices, MoIP can be classified as softswitch-controlled MoIP and auto-switching MoIP.

Softswitch-Controlled MoIP

The basic flow of the softswitch-controlled MoIP service is as follows:

1. Establish a call. If the MoIP service is configured on the softswitch, the softswitch sends a command to the MG instructing the MG to detect the modem event.
2. The calling party and called party start communicating with each other.

3. During the call, when the MG detects the ANS or ANSAM modem start event (low-speed modem signal), or detects the ANSBAR or ANSAMBAR modem start event (high-speed modem signal), the MG sends the event to the softswitch.
4. According to the event, the softswitch sends a command instructing the MG to switch the DSP channel of the calling and called parties to the low-speed or high-speed modem mode.
5. According to the command sent by the softswitch, the MG switches the DSP channel to the corresponding modem mode. At this stage, the MG adopts the encoding format and port number specified by the softswitch.
6. The settings of echo cancellation (EC), voice activity detection (VAD), and DSP working mode are as follows:
   a. Low-speed modem: EC-ON, VAD-OFF, DSP working mode-modem mode
   b. High-speed modem: EC-OFF, VAD-OFF, DSP working mode-modem mode
7. After the modem data is transmitted, if the conversation proceeds, the DSP working mode does not automatically switch from the modem mode to the voice mode, because the modem end event is not issued. As a result, the quality of the voice service may be affected.

Auto-Switching MoIP

The basic flow of the auto-switching MoIP service is as follows:

1. Set up a conversation.
2. The MGs at both ends check for the modem event on the IP side and the TDM side. When the modem event is detected, if the modem transmission mode is configured as auto-switching, the coding mode is switched to G.711 (the a/μ law is configurable), and the DSP parameters are modified according to the modem mode (high-speed/low-speed) detected.
3. When the modem service is terminated, the call is released.

MGCP-Based FoIP

FoIP refers to the fax service provided on the IP network or between the IP network and the traditional PSTN network. The fax machine can be regarded as a special modem. In the FoIP negotiation, the modem negotiation is performed before the fax negotiation.

According to the transmission protocol adopted, there are two modes of fax services carried on the IP network: the T.30 transparent transmission mode and the T.38 mode. According to different control devices, FoIP can be classified as softswitch-controlled FoIP and auto-switching FoIP.

Softswitch-Controlled FoIP

The fax service can be classified into high-speed fax and low-speed fax. The softswitch-controlled low-speed fax service supports the T.30 transparent transmission mode and the T.38 mode. The basic service flow is as follows:

1. Configure the fax service and fax flow on the MGs and the softswitch.
2. After the voice channel is set up, the softswitch instructs the MG to detect the fax event and modem event.
3. When detecting the fax event, the MG reports the event to the softswitch. The event can be a low-speed modem event (ANS or ANSAM) or a low-speed fax event (V.21Flag).

4. According to the preset fax flow, the softswitch instructs the MGs at both ends to change the DSP channel working mode to the T.30 transparent transmission mode or T.38 mode.
5. The fax starts.
6. After the fax is complete, if the MG detects the fax end event, the MG reports the event to the softswitch.
7. The softswitch instructs the MGs at both ends to change the DSP channel working mode to the voice mode.
8. The voice service proceeds.

The softswitch-controlled high-speed fax service supports the T.30 transparent transmission mode. The basic service flow is as follows:

1. Configure the fax service and fax flow on the MGs and the softswitch.
2. After the voice channel is set up, the softswitch instructs the MG to detect the fax event and modem event.
3. When detecting a fax event, the MG reports the event to the softswitch. The event can be a high-speed modem event (ANSBAR or ANSAMBAR) or a low-speed fax event (V.21Flag; if the peer end is a low-speed fax machine or the network quality is poor, the fax speed is automatically decreased and this event is reported).
4. According to the preset fax flow, the softswitch instructs the MGs at both ends to change the DSP channel working mode to T.30 transparent transmission mode.
5. The fax starts.
6. After the fax is complete, if the MG detects the fax end event, the MG reports the event to the softswitch.
7. The softswitch instructs the MGs at both ends to change the DSP channel working mode to the voice mode. The voice service proceeds.

Auto-Switching FoIP

The auto-switching fax service supports the T.30 transparent transmission mode and the T.38 mode. The basic service flow is as follows:

1. Configure the auto-switching fax service on the MGs at both ends.
2. Establish a call and use the voice service.
3. The MG checks for the fax event on the IP side and the TDM side. When detecting the fax event, the MG changes the DSP channel working mode to the T.30 transparent transmission mode or the T.38 mode.
4. After the fax is complete, when the MG detects the fax end event, the MG changes the DSP channel working mode to the voice mode.
5. The voice service proceeds.

Common Fax Protocols

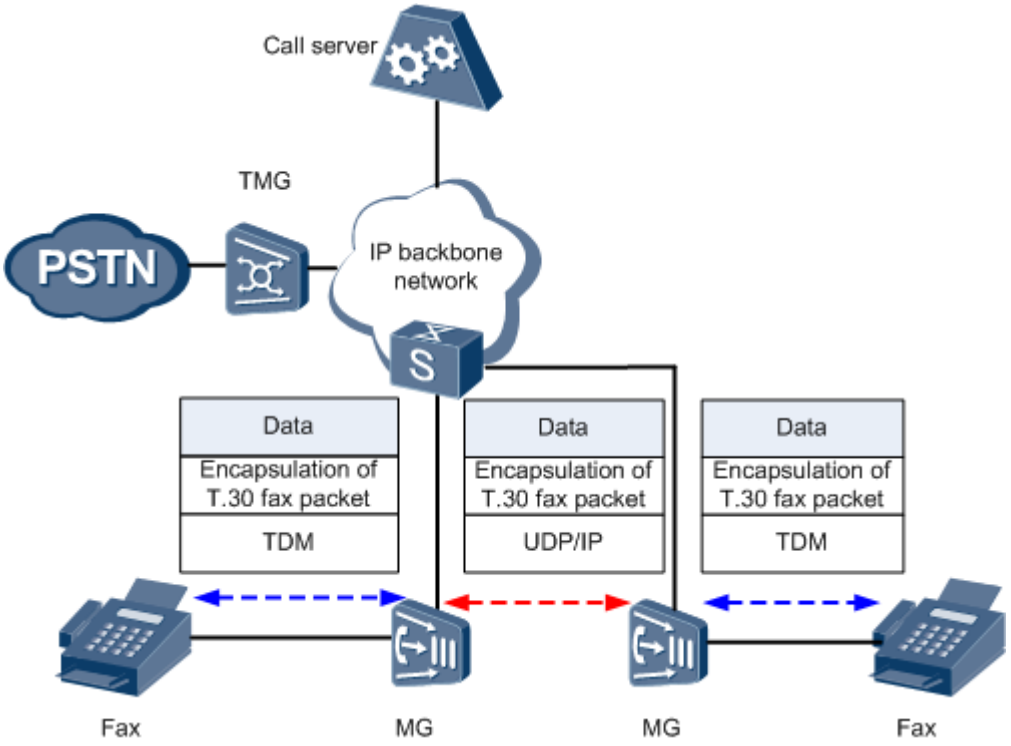
Two protocols are usually used for implementing the fax service on the packet voice network: the ITU-T Recommendation T.30 and ITU-T Recommendation T.38.

T.30 is based on the PSTN network. T.30 particularly defines the flow for transmitting fax signals on the PSTN network. It also defines the modulation mode (V.17/V.21/V.27/V.29/V.34) and transmission format (HDLC) of data, and the physical standard for fax signals. The T.30 fax messages and data can be transmitted transparently between MGs. This is called the T.30

transparent transmission mode. The quality of the fax in this mode may not be high due to packet loss, latency, and disorder on the IP network.

T.38 is a real-time fax mode based on the IP network. In this mode, the MG terminates the T.30 signals sent from the fax machine, and transmits the data to the peer MG in the T.38 mode. The peer MG then receives the T.38 packets and converts the packets into T.30 signals. The merit of the T.38 fax is that the fax packets have a redundancy processing mechanism and do not strictly rely on the quality of the network (the fax service can be processed even when a 20% packet loss occurs on the network). The demerit is that the DSP chip needs to participate in parsing the T.30 signals. Because there are various types of terminals on the network, the compatibility problem may arise.Figure 23-53 illustrates the principle of the T.38 fax.

Figure 23-53 Principle of the T.38 fax



23.6.3 MGCP Standards and Protocols Compliance

- RFC2705
- RFC3405
- T.30: It is based on the PSTN network. T.30 particularly defines the flow for transmitting fax signals on the PSTN network. It also defines the modulation mode (V.17/V.21/V.27/V.29/V.34) and transmission format (HDLC) of data, and the physical standard for fax signals. The T.30 fax messages and data can be transmitted transparently between MGs. This is called the T.30 transparent transmission mode. The quality of the fax in this mode may not be high due to packet loss, latency, and disorder on the IP network.
- T.38: It is a real-time fax mode based on the IP network. In this mode, the MG terminates the T.30 signals sent from the fax machine, and transmits the data to the peer MG in the T.38 mode. The peer MG then receives the T.38 packets and converts the packets into

T.30 signals. The merit of the T.38 fax is that the fax packets have a redundancy processing mechanism and do not strictly rely on the quality of the network (the fax service can be processed even when a 20% packet loss occurs on the network). The demerit is that the DSP chip needs to participate in parsing the T.30 signals. Because there are various types of terminals on the network, the compatibility problem may arise.

23.7 H.248 Voice Feature

This topic first describes the H.248 protocol, and then describes the protocol mechanism, and last describes the application of H.248 in VoIP, MoIP, and FoIP.

23.7.1 Introduction to the H.248 Feature

Definition

H.248 is a media gateway control protocol through which the media gateway controller (MGC) controls the media gateway (MG) so that interoperability is implemented between different media. ITU-T issued the first version of this protocol in June 2000.

Purpose

Compared with MGCP, H.248 has the following merits:

- Supports more types of access technologies, and is more thorough and complete in standardization
- Compensates for the deficiency of MGCP in descriptiveness, is applicable to larger networks and has better extensibility and flexibility
- Carried on various protocols, such as UDP/SCTP (MGCP is carried on UDP)

23.7.2 H.248 Principles

Mechanism of the H.248 Protocol

Termination ID

A termination ID identifies a termination that is going to register or deregister a service. The termination ID of each termination is unique. During service configuration, the termination ID corresponding to each termination must be configured on the MG and the MGC. The root termination ID represents an entire MG. The ServiceChange command executed on the root termination ID is effective on an entire MG. The wildcarding principle is that the ALL wildcard (*) can be used but the CHOOSE wildcard ($) cannot be used.

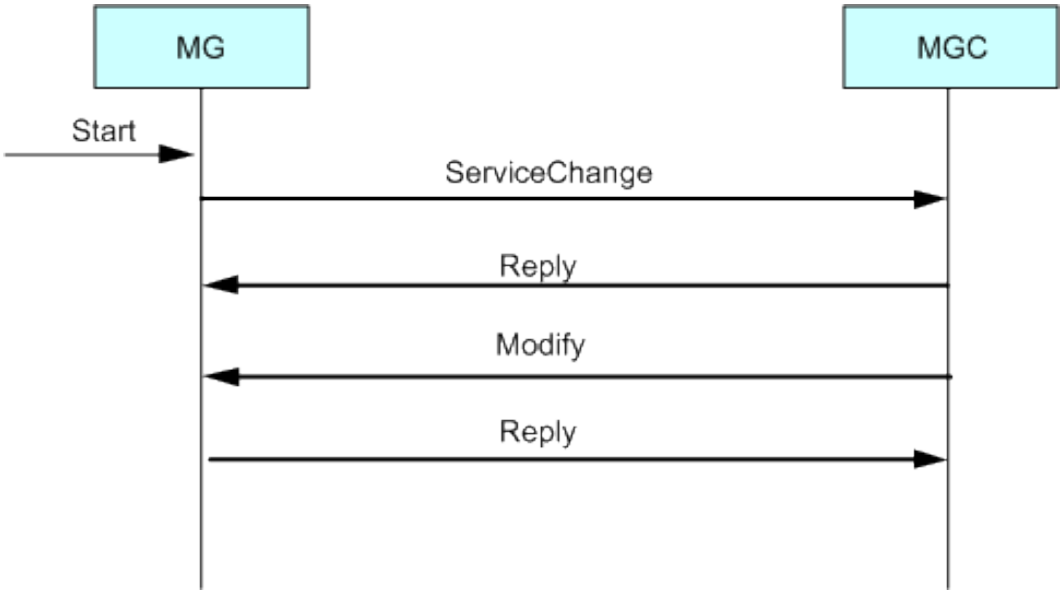
Registration Mechanism of the H.248 Interface

The MG sends the ServiceChangeRequest command to inform the MGC that a user or a group of users are about to register or deregister service. After this command is executed successfully, the termination status is changed to InService or OutOfService. In addition, the MGC can unsolicitedly send the ServiceChangeRequest command to request the MG to register or deregister service for a user or a group of users.

NOTE
Currently, the MG does not support the MGC to unsolicitedly send the ServiceChangeRequest command requesting the MG to register service for a user or a group of users.

Figure 23-54 shows the registration flow of the MG.

Figure 23-54 Registration flow of the MG



Description of the flow:

1. The MG sends the ServiceChangeRequest command to the MGC. In the command, TerminationId is Root, Method is Restart, and ServiceChangeReason is 901 (cold boot, registering for the first time after power-on), 902 (warm boot, through command lines), or 900 (in other cases).
2. The MGC sends the Reply message to the MG indicating the successful registration.
3. The MGC sends the Modify command to the MG requesting the MG to detect the offhook of all users (al/of).
4. The MG responds to the MGC with the Reply message.

Heartbeat Mechanism of the H.248 Interface

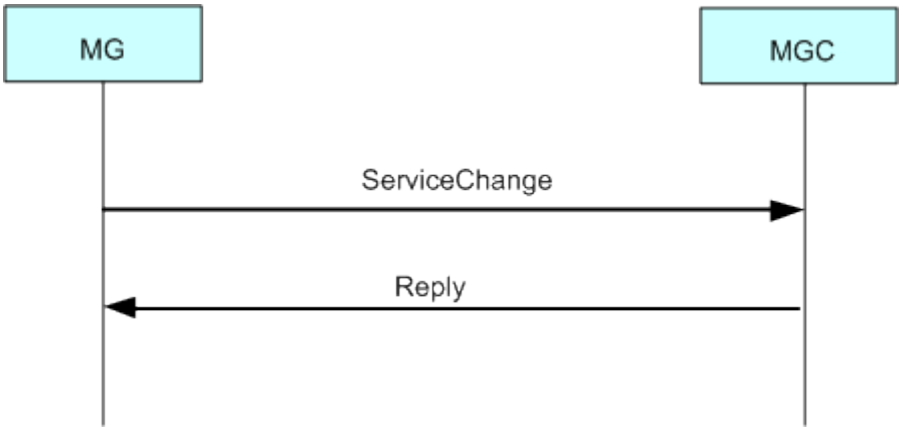
After the registration is successful, the MG and the MGC maintain communication by sending each other the heartbeat message Notify (it/ito). By default, the heartbeat message is sent every 60s. The sending interval can be set within the range of 5-655s.

After the MG sends the first heartbeat message to the MGC, if the MG does not receive the heartbeat response from the MGC before the preset interface heartbeat timer (for example, the length of three sending intervals) times out, the MG sets the interface status to "wait for response". Then, the MG keeps initiating a registration with the MGC. If dual-homing is configured, the MG initiates registration with the two MGCs alternatively. The registration is initiated once every 30s, every three trials of registration are one round, and every registration message is re-transmitted 7 times. Therefore, 24 registration messages in total are transmitted within 90s. Then, the MG starts the next round of registration with the other MGC.

Deregistration Mechanism of the H.248 Interface

Figure 23-55 shows the unsolicited deregistration flow of the MG.

Figure 23-55 Unsolicited deregistration flow of the MG

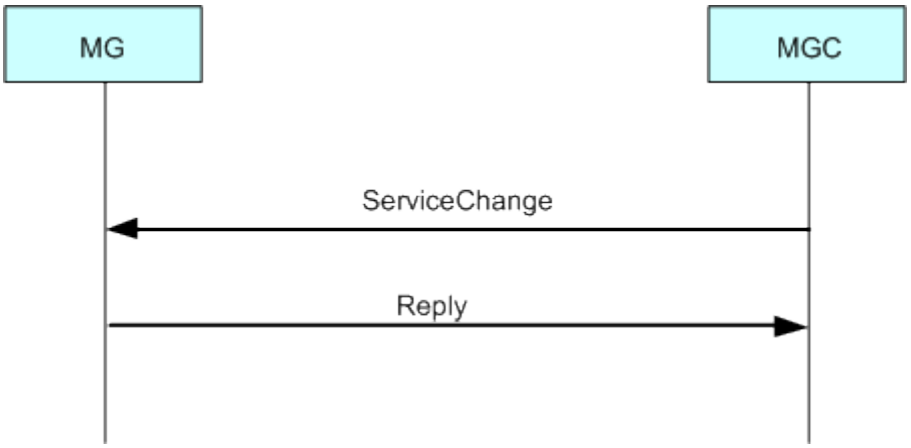


Description of the flow:

1. The MG sends the ServiceChangeRequest command to the MGC. In the command, TerminationId is Root, Method is Forced, and ServiceChangeReason is 905 ("905" indicates that the termination is taken out of service because of maintenance operation, and now the MG uses "905" to initiate a deregistration request through command lines).
2. The MGC sends the Reply message to the MG indicating a successful deregistration.

Figure 23-56 shows the flow of the MGC unsolicitedly deregistering the MG.

Figure 23-56 Unsolicited deregistration flow of the MGC



Description of the flow:

1. The MGC sends the ServiceChangeRequest command to the MG. In the command, TerminationId is Root, Method is Forced, and ServiceChangeReason is 905.
2. The MG responds to the MGC with the Reply message. The MA5600T/MA5603T/MA5608T (MG) supports the registration and deregistration of not

only an entire MG but also a single termination. The service status of a single user can be changed through the registration and deregistration of a single termination.

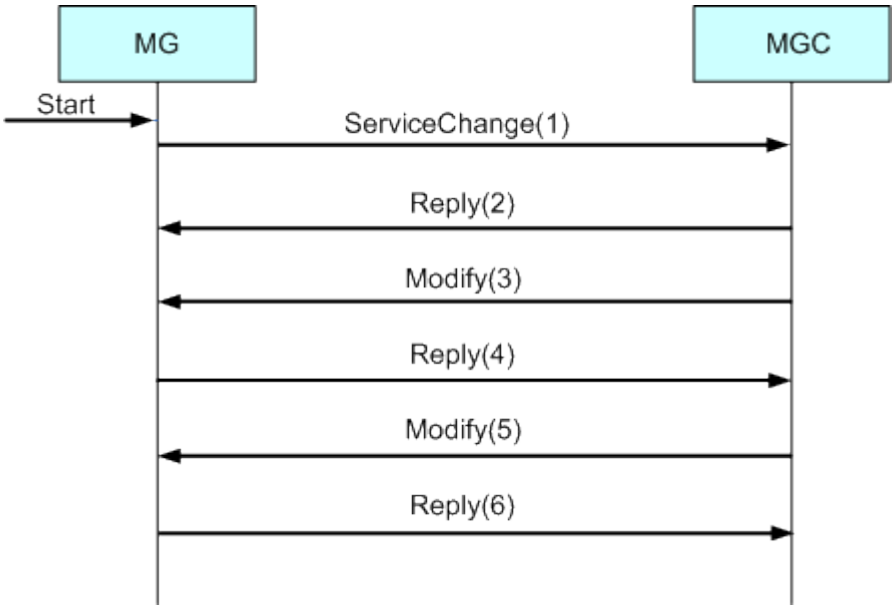
Authentication Mechanism of the H.248 Interface

Authentication is a security mechanism through which the MGC authenticates the legality of the MG user. The purpose of authentication is to prevent unauthorized entities from establishing illegal calls or interfering with legal calls through the H.248 or MGCP protocol. Authentication can be implemented only when it is also supported by the softswitch interconnected with the MG.

- In H.248, the implementation of authentication complies with RFC2402.
- MD5 is adopted as the encryption algorithm.

Figure 23-57 shows the authentication flow.

Figure 23-57 Authentication flow



The basic flow is as follows:

1. The MG sends the ServiceChange command to register with the MGC. The command contains the digital signature of the MG.
2. After receiving the ServiceChange command, the softswitch verifies the MG and sends a reply.
3. The softswitch sends the Modify message to the MG. The message contains the required algorithm ID and random number.
4. The MG verifies the message sent by the softswitch and sends a reply.
5. The softswitch authenticates the MG periodically.
6. The MG sends replies to the softswitch.

H.248-Based VoIP

Figure 23-58 illustrates the principle of the call establishment and release in the H.248-based VoIP service.

Figure 23-58 Principle of the VoIP feature that supports the H.248 protocol

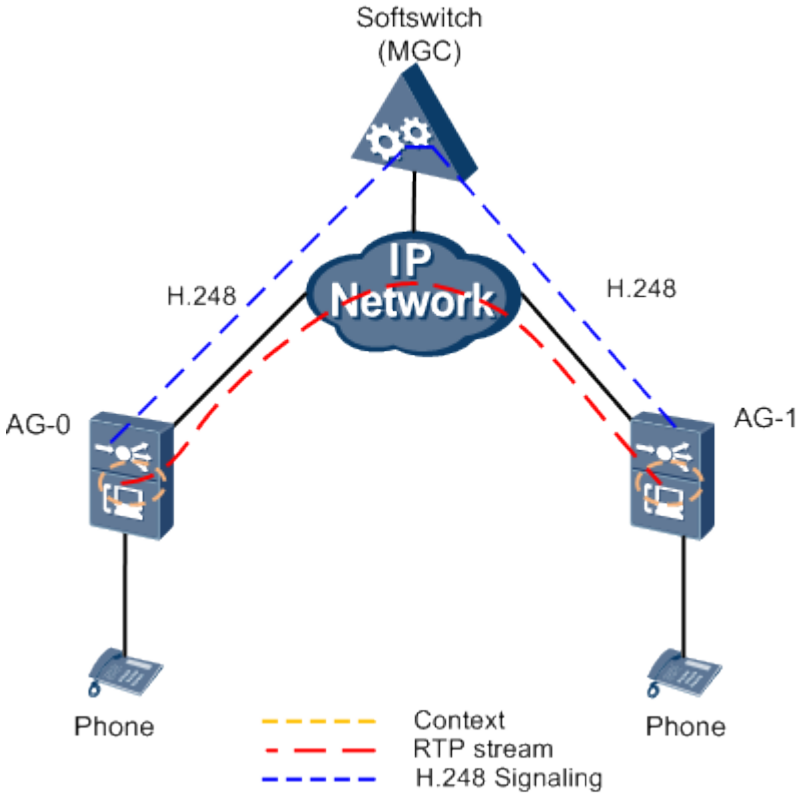
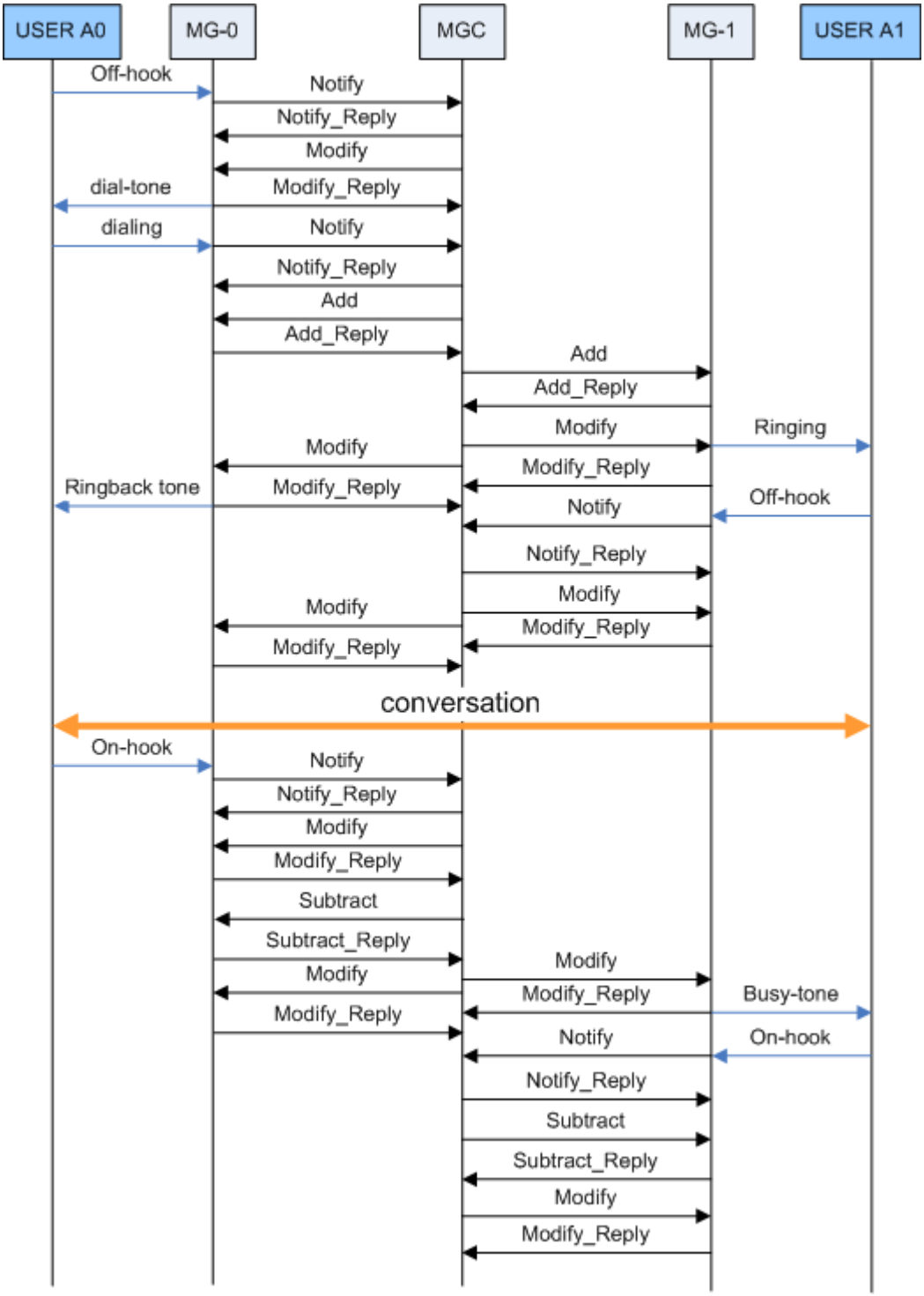


Figure 23-59 illustrates the basic flow of a call establishment and release.

Figure 23-59 H.248-based call flow



1. MG-0 detects the offhook of user A0, and notifies the MGC of the offhook event through the Notify command.
2. After receiving the offhook event, the MGC sends a digitmap to MG-0, requests MG-0 to play the dial tone to user A0, and at the same time checks for the digit collection event.

3. User A0 dials a telephone number, and MG-0 collects the digits according to the digitmap issued by the MGC. Then, MG-0 reports the result of digit collection to the MGC.
4. The MGC sends the Add command to MG-0 for creating a context and adding the termination and RTP termination of user A0 into the context.
5. After creating the context, MG-0 responds to the MGC. The response contains the session description that provides the necessary information for the peer end to send the packet to MG-0, such as the IP address and UDP port number.
6. The MGC sends the Add command to MG-1 for creating a context and adding the termination and RTP termination of user A1 into the context, and then issues the IP address/UDP port number of user A0 to user A1.
7. After creating the context, MG-1 responds to the MGC. The response contains the session description that provides the necessary information for the peer end to send the packet to MG-1, such as the IP address and UDP port number.
8. MG-1 detects the offhook of user A1, and then reports the offhook event to the MGC. The softswitch (MGC) sends the Modify command to stop the ring back tone of user A0 and the ringing of user A1.
9. The MGC sends the session description of MG-1 to user A0 through the Modify command. Then, the conversation is set up between users A0 and A1.
10. MG-0 detects the onhook of user A0, and notifies the MGC of the onhook event through the Notify command.
11. The MGC sends the Modify command to MG-0 and MG-1 respectively to modify the RTP mode to receive-only.
12. The MGC sends the Modify command to MG-1 requesting MG-1 to play the busy tone to user A1, and at the same time checks for the onhook event.
13. The MGC sends the Subtract command to MG-0, requesting MG-0 to release the resources that are occupied by the call of user A0.
14. MG-1 detects the onhook of user A1, and notifies the MGC of the onhook event through the Notify command.
15. The MGC sends the Subtract command to MG-1, requesting MG-1 to release the resources that are occupied by the call of user A1.
16. The call between users A0 and A1 is terminated, and all the resources occupied by the call are released.

H.248-Based MoIP

H.248 is similar to MGCP; therefore, for the core flow of the connection establishment and release of the H.248-based MoIP service, see MGCP-Based MoIP.

H.248-Based FoIP

H.248 is similar to MGCP; therefore, for the core flow of the H.248-based fax, see MGCP-Based FoIP.

23.7.3 H.248 Standards and Protocols Compliance

- RFC3525

23.8 POTS Access

This topic describes the features in relation to the POTS interface, including basic features such as ringing and Z interface and enhanced features.

23.8.1 Introduction to POTS Access

Definition

Plain Old Telephone Service (POTS) is the traditional basic telephony service provided using twisted-pair copper lines. POTS Voice interface features are the features implemented on the voice interface.

Purpose

The purpose is to provide the standard-compliant voice interface that has the reliable protection capability and intelligent energy-saving function.

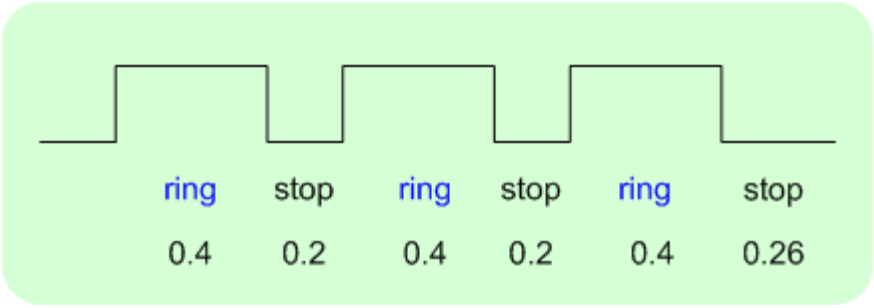
23.8.2 Ringing

A subscriber's phone rings, indicating that an incoming call is waiting on the line. The POTS board plays the ring tone to plain old telephone service (POTS) subscribers, and terminals play the ring tone to integrated services digital network (ISDN) subscribers. This section describes the ring tone played by the POTS board.

Break-Make Ratio of the Ring Tone Played by the POTS Board

The break-make ratio varies depending on countries. For example, the break-make ratio in China is 1:4, that is, the ring tone is played in the cycle of one-second ringing followed by four-second mute. Figure 23-60 shows a six-segment ring with break-make ratio 0.4:0.2:0.4:0.2:0.4:2.6, which cycles in the following mode: ringing for 0.4 seconds and mute for 0.2 seconds twice followed by 0.4-second ringing and 2.6-second mute.

Figure 23-60 Ring example



Application

In some countries, break-make ratio is different for different services. For example, break-make ratio 1:4 is used for local calls, 1:2 for toll calls, and 0.4:0.2:0.4:4 for intra-Centrex calls.

23.8.3 POTS Interface Protection

Introduction

Service boards are connected to user terminals through subscriber cables. Some subscriber cables may be routed under the ground or overhead and some may be routed in parallel with the mains AC power cables. In these cases, a high voltage may be generated on subscriber cables because of the lightning attack, contact with power lines, and induction of power lines. The high voltage may damage the ports on service boards. Therefore, service boards must be equipped with the protection capability to prevent occurrence of the preceding problems.

23.8.4 Features of the POTS Line Interface

Standards of the POTS Line Interface

Major standards of the voice line interface are as follows:

- ITU-Q552: It defines transmission specifications of the Z interface.
- ES 201970: It defines basic hardware features of the voice interface.
- YD751 - Network Entry Checking Methods for Telephone Exchange Equipment: It describes the voice interface standards defined by China.

The voice technologies have gone through a long period of development. Almost every country has its own standards. The preceding standards are related to the basic features of boards. Special features must be tailored to meet requirements of different countries. For example, British Telecom (BT) has its own standards, namelyBTNR315. Some of the requirements in the BTNR315 are very special and can be met only by special boards.

Basic Features of the Z Interface

Basic features of the Z interface supported by the voice interface board of the MA5600T&MA5603T&MA5608T are summarized as follows:

- Battery feeding (B)
  - Batter feeding refers to the supply of the voltage and the current to terminals (such as telephones) to ensure the normal operation of terminals.
  - When the telephone is in the on-hook state, the voltage of the board of the MA5600T&MA5603T&MA5608T is generally 48 V. When the telephone is in the off-hook state, the board of the MA5600T&MA5603T&MA5608T supports the constant-current feeding at 20 mA, 25 mA, or 30 mA. The feeding current can be configured according to the actual requirement.
  - The off-hook feeding of the port can be automatically adjusted. If the length of the loop is short, the port is fed with the constant current. If the length of the loop is long, the port automatically adjusts the loop current based on the preset threshold. This design ensures the compliance with the related standards and optimizes the power consumption of the port.
  - If the feeding current is 25 mA and the voltage is -48 V, the feeding current is equal to or larger than 25 mA when the loop resistance is less than 1200 ohm, and the feeding current is larger than 18 mA when the loop resistance is 1800 ohm.
  - The boards of the MA5600T&MA5603T&MA5608T also support the 40-mA feeding current. The 40-mA feeding current increases the power consumption of ports and thus is not recommended. If the 40-mA feeding current is configured, the

number of ports configured with the 40-mA feeding current cannot exceed five on each board.

- Ringing (R)
  - Ringing refers to the supply of the ring current to telephones so that telephones can ring to inform subscribers of incoming calls. The boards of the MA5600T&MA5603T&MA5608T are designed with the balanced ringing feature.
  - The concept of the balanced ringing is put forward based on the concept of the traditional imbalanced ringing. The traditional imbalanced ringing is classified into two types: (1) In a subscriber line, A line is 0 V and B line is -48 V DC overlaid with the 75 Vrms AC signals. (2) In the subscriber line, A line is -48 V and B line carries the 75 Vrms AC signals. In the case of the balanced ringing, both A and B lines of the subscriber line have the AC signals. The AC signals of the A and B lines are of the same frequency and opposite phases, that is, differential signals. The frequency of signals in the case of the balanced ringing can be set to 16 Hz, 25 Hz, or 50 Hz.
  - The amplitude of the ringing current can reach up to 70 Vrms. The amplitude of the ringing current on a terminal can exceed 35 Vrms if the line impedance is 1400 ohm (5-km lines with the core diameter of 0.4 mm) and the terminal impedance is 4000 ohm. The amplitude of the 50-Vrms ringing current is configurable. This configuration is mainly applicable to the short loop with a length less than one kilometer, aiming to substantially reduce the power consumption of ringing on the ports.
  - The DC offset provided by boards can reach 20 V, which ensures reliable ringing when the distance is long.
  - The break-make ratio of the ringing current can be configured to meet requirements of different carriers in the world.
- Over-voltage protection (O)

  Over-voltage protection is one of the interface protection measures.
- Supervision (S)

  Supervision refers to the detection of telephone state, such as on-hook, off-hook, and off-hook in the ringing state. The terminal state can be learned through detection. The terminal state detection is the basis of some calls.
- Code/Decode (C)

  Coding/Decoding refers to the process that analog signals of the subscriber line are converted into digital signals and compressed according to the A/U law.
- H - Hybrid circuit

  Hybrid circuit refers to the conversion from the 2-wire analog interface to the 4-wire digital interface on the board and implementation of the balanced matching with the impedance of the subscriber line.
- Test (T)

  For details about the test function, see 23.15 Voice Service Maintenance and Diagnosis.

Interface Impedance, Transmission Specifications, and Gain

The voice interface board of the MA5600T&MA5603T&MA5608T supports the configuration of the interface impedance and gain.

At present, sixteen common interface impedances can be configured, see as the Table 23-14:

Table 23-14 Lists of common interface impedances

ID	Impedance of port and usage description
0	(200+680|| 100nf): bureau machine in China
1	(200+560|| 100nf): user machine in China
2	600ohm: a common interface
3	(150+510|| 47nf): interface of Russian
4	(220+820|| 115nf): widely used in countries like Germany
5	(220+820|| 120nf): widely used in Germany
6	900ohm: seldom used
7	(800|| 50nf): interface of Brazil
8	(Zin=87+1052||228nF+229||28.4nF, Zload=93+615|| 471nF+179||495nF+244||32nF): interface of BT0
9	(Zin=370+620|| 310nf,Zload=600): interface of HK_BT3
10	(Zin=270+264|| 357nf+1434|| 265nf,Zload=600): interface of HK_BT5
11	(BT0 without AGC): interface of BT1
12	(Zin=87+1052||228nF+229||28.4nF, Zload=270+264|| 357nF+1434||265nF): Interface of BT2
13	(Zin=87+1052||228nF+229||28.4nF, Zload=164+162|| 363nF+1227||350nF): interface of BT3
14	(Zin= 270+750|| 150nf): a common interface widely used in Europe
15	(Zin= 370+620|| 310nf): interface of New Zealand

The interface transmission gain is also configurable. The send gain is generally in the range of +4 dB and -6 dB and the receive gain is in the range of 0 dB to -12 dB. The gain can be configured at the step of 0.5 dB.

The transmission specifications of the boards are fully compliant with the ITU-Q522 test requirements. If the interface impedance is not one of the preceding eight types, independent software can be developed to support the interface impedance.

Digit Collection

The voice interface board of the MA5600T&MA5603T&MA5608T supports the pulse-based digit collection.

Old-fashioned telephones generally adopt the pulse dialing mode, while new telephones adopt the DTMF dialing mode. Most telephones support the pulse dialing mode.

The service boards support the pulse-based digit collection at the speed of 8 pps to 12 pps. The break-make ratio is in the range of 50% and 80%. The interval of pulses is configurable and is in the range of 100 ms and 2 s. The default interval of pulses is 300 ms.

The DTMF digit collection is completed by the DSP instead of the service boards.

Charging Signals

Service boards support three charging modes, namely polarity reverse, 12/16KC, and counter impulse delivery.

- Polarity reverse: The voltage polarity between A and B lines of the subscriber line is reversed. Some terminals detect this type of reverse for charging purpose.
- 12/16KC: The service board sends the 12000 Hz/16000 Hz sine AC signals at a specific interval to the terminals.
- Counter impulse delivery: The service board sends pulse signals to the terminals. Charging is implemented based on the pulse signals.

All ports of the service board support both the fast and slow polarity reverse features. Fast polarity reverse is generally completed within 3 ms, which meets the time requirements of polarity reverse of some telephones. Slow polarity reverse is generally completed within 80 ms. It can substantially reduce the interference to the line during the polarity reverse and is compatible with the DSL transmission on the same line.

The service board supports the 12/16KC charging. In the 12/16KC charging mode, the amplitude of the 12/16KC signals is configurable. The amplitude can be set to 0.45 Vrms, 0.775 Vrms, 1 Vrms, 1.5 Vrms, 2 Vrms, or 2.5 Vrms. The maximum value is 2.5 Vrms (200 ohm). In addition, the break-make ratio of KC signals is also configurable. By default, the Make duration is 100 ms and the Break duration is 300 ms. Both the Make duration and the Break duration range from 10 ms to 500 ms.

The service board also supports the counter impulse delivery charging. Some attributes of this charging mode, such as the pulse width and number of pulses sent per minute, are configurable.

Current Reduction of Locked Ports

When a phone connected to a port is in off-hook state for a long time but the conversation is not going on, the service board can lower the current of the port to less than 12 mA to reduce the power consumption of the port.

Short Loop Feeding

When the length of the line is short, the service board uses the low voltage for feeding to reduce the power of the port. When the length of the line becomes long, the service board automatically uses the voltage higher than the previous low voltage to meet the application requirement.

Power Cut-off

Feeding of ports that are not allocated with numbers can be cut off to reduce the power consumption of the ports.

On-Hook Transmission

Service boards support the on-hook and off-hook transmission functions, such as the caller identification display service and the fixed network short message service.

Ringer Equivalence Number

Ringer equivalence number (REN) refers to the number of telephones that can be connected to the same port.

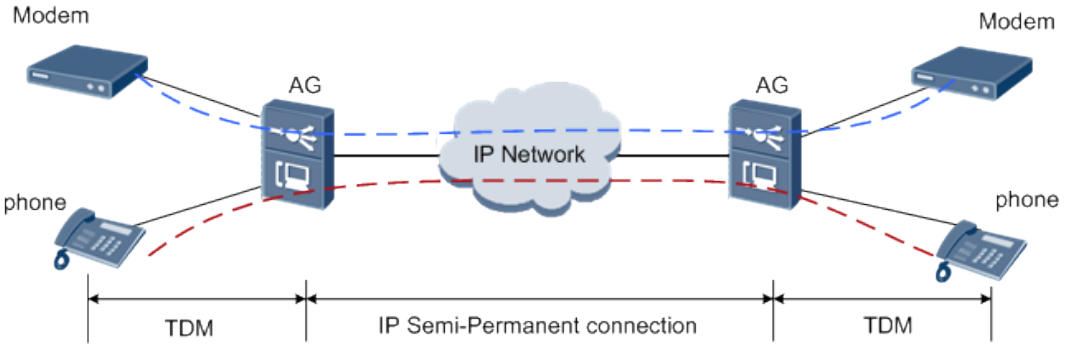
23.8.5 POTS IP SPC

The semi-permanent connection (SPC) exclusively occupies a constant voice channel to meet the communication requirements and ensure the communication quality for special and key access users. To configure an IP SPC, configure the data (including the local IP address, local UDP port, remote IP address, and remote UDP port), set up an IP direct connection between the two ends of the voice service. In this manner, the voice media data can be directly transmitted to the peer end.

Networking Description

Figure 23-61 shows the networking of a POTS IP SPC. An IP SPC is set up between the POTS port on AG 1 and POTS port on AG 2. Users under these two POTS ports are permanently online and can communicate with each other without dialing a number.

Figure 23-61 Networking of a POTS IP SPC



Application Scenarios

A POTS IP SPC is mainly used in the following two scenarios:

- Special modems are connected to a POTS port and these modems require direct communication without dialing and require online permanently.
- Two phones are online permanently without dialing. This scenario is rare, which may be applied to some special dispatching phones.

23.9 ISDN Access

The integrated services digital network (ISDN) is a CCITT standard, providing integrated transmission service for voice, video, and data. The ISDN enables the voice, video, and data to be transmitted on the data channel simultaneously.

23.9.1 Introduction to ISDN

Definition

The integrated services digital network (ISDN) is a CCITT standard, providing integrated transmission service for voice, video, and data. The ISDN enables the voice, video, and data to be transmitted on the data channel simultaneously.

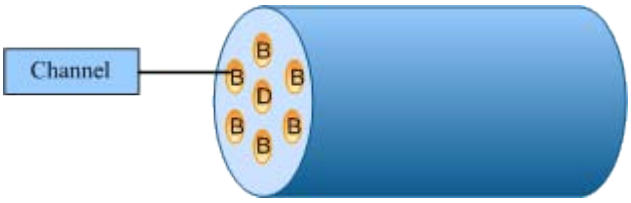
The ISDN supports two types of services:

- Basic rate interface (BRI): provides a rate of 144 kbit/s, including two B channels and one D channel. The rates of B and D channels are 64 kbit/s and 16 kbit/s, respectively.
- Primary rate interface (PRI): provides a rate of 2.048 Mbit/s, including 30 B channels and one D channel. The rates of both B and D channels are 64 kbit/s.

ISDN networks support B and D channels, shown in Figure 23-62.

- B channels are used for carrying services.
- D channels are used for transmitting call control signaling as well as maintenance and management signaling.

Figure 23-62 ISDN channels



Purpose

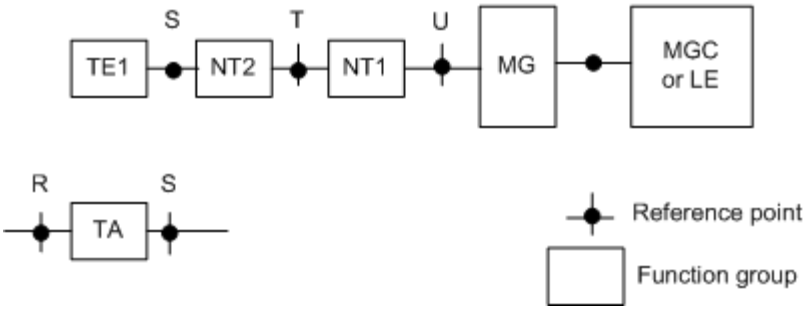
The ISDN access on the media gateway provides integrated transmission services, such as voice, video, and data for the users.

23.9.2 ISDN Protocol Model

ISDN Reference Model

Figure 23-63 shows the ISDN reference model.

Figure 23-63 ISDN reference model



NOTE

- Network Termination 1 (NT1) operates at Layer 1 (physical layer) of the OSI model and implements the physical and electronic specifications into the ISDN network.
- Network Termination 2 (NT2), an intelligent device, functions as a terminal control device, such as a private automatic branch exchange (PABX) or a LAN router, and operates at Layer 2 and 3 of the OSI model.
- Terminal Equipment Type 1 (TE1) is the standard ISDN device having the standard S port, such as an ISDN phone and G4 fax machine. It can be directly connected to NT2 or NT1.
- Terminal Equipment Type 2 (TE2) is a non-standard ISDN device, such as a PC or an X.25 packet terminal. It cannot be directly connected to NT2 or NT1 but must be connected to the S port through the TA.
- Terminal Adapter (TA) connects a non-ISDN terminal (TE2) to the user-network interface (UNI) of the ISDN.
- U reference point, also called the U port, is the line interface locates between the ISDN BRA network and user. Digital signals are transmitted through twisted pairs through the coding (such as 2B1Q coding) defined by the U port.
- S reference point, also called the S port, is the line interface locates between the ISDN terminal (TE1 or TA) and NT.
- T reference point locates between NT1 and NT2. If there is no NT2, S referent point and T reference points are combined as S/T reference point, also called the S/T port. It uses 4–wire for transmission, such as a common network cable.
- R reference point locates between the TA and TE2 (non-ISDN standard device) and provides interfaces (the RS-232 interface for PCs and X.25 interface for X.25 devices) that allow the non-ISDN standard device to access the ISDN.

The ISDN user accesses the MA5600T/MA5603T/MA5608T through the U reference point. The actual terminal on the user side may support NT 1, NT 2, and TE 1 functions at the same time. When VoIP is used for upstream transmission, the IUA protocol is used to load the Q.931 call signaling of the ISDN between the MG and MGC, and the H.248 protocol or MGCP signaling is used to control the media connection on the MG.

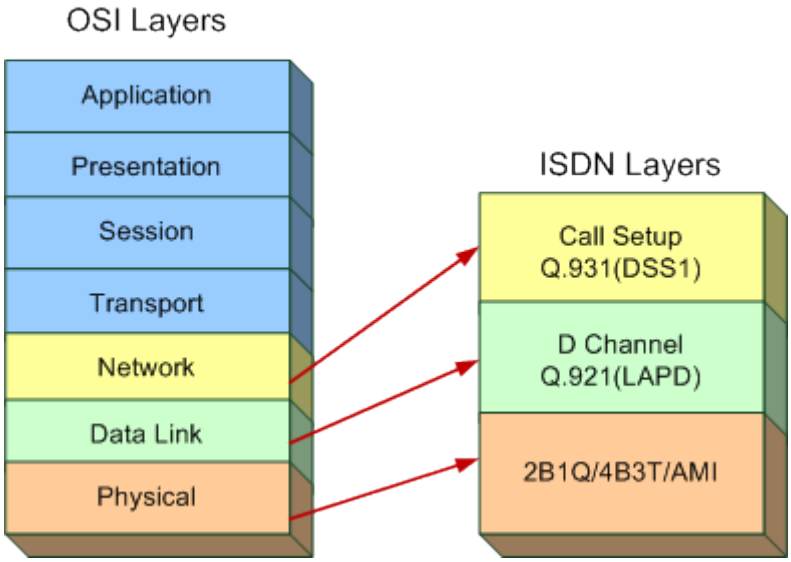
ISDN Protocol Stack Model

Figure 23-64 shows the mapping relationship between the ISDN protocol and OSI model. Layers in the ISDN protocol stack map the physical layer, data link layer, and network layer in the following OSI model.

- ISDN physical layer: For users, the ISDN physical layer is on S reference point or T reference point. This layer has the following major functions: coding, full-duplex transmission, channel multiplexing, port activation and depolarization, feeding, and termination identification. This layer can multiplex multiple links at the data link layer and use AMI, 4B3T and 2B1Q for coding.

- ISDN data link layer: ISDN does not define Layer 2 protocols dedicated to B channels. Any Layer 2 protocols can be used between two communicating devices after negotiation as long as they can transparently transmit data on B channels. The link access procedure on the D channel (LAPD) protocol defined in Q.921 (a reliable transport protocol) is used for D channels, which is mainly used to carry messages and data generated by Layer 3 entities.
- ISDN network layer: ISDN does not define Layer 3 protocols dedicated to B channels. Layer 3 protocol Q.931 for D channels is mainly used to control and manage connection setup and release on B channels.

Figure 23-64 Mapping relationship between the ISDN protocol stack model and OSI model



ISDN Protocol Processing Model

Figure 23-65 and Figure 23-66 show the ISDN protocol processing model. ISDN involves the following protocols:

- Q.921: Defines LADP. It is a reliable transport protocol.
- Q.931: Defines the procedure of processing and controlling messages and state machines that are used for calls (including circuit switching calls and packet switching calls) between the user-side device and network-side device.
- SCTP: A transport protocol in the SIGTRAN protocol stack, which is a reliable transport protocol on top of protocols (such as IP) providing unreliable transmission services. SCTP transmits acknowledged, error-free, and repetition-free data and ensures real time transmission to some extent.
- IUA: ISDN Q.921 user adaption layer protocol

Figure 23-65 H.248 protocol processing model

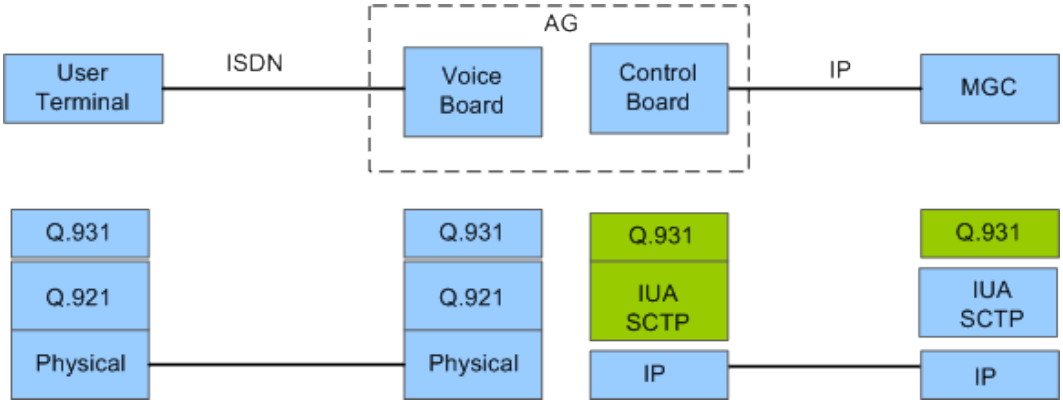
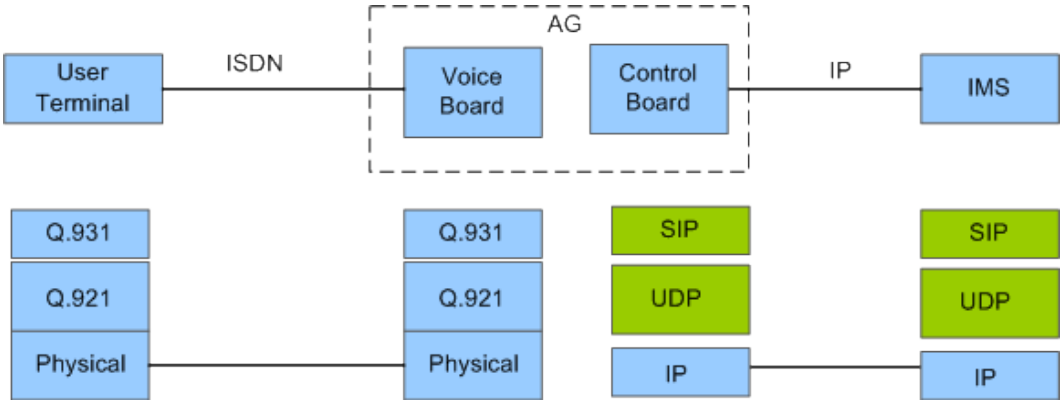


Figure 23-66 SIP protocol processing model



Exchange between protocols

For the H.248 protocol:

- Data transmission from ISDN user terminals to voice boards uses Q.921 and Q.931.
- Voice boards terminate Q.921 messages and send Q.931 messages to the CPU of the control board using the master-slave serial port communications protocol. Then the CPU of the control board uses IUA to packetize Q.931 signaling carried on SCTP links and sends the packed signaling to the MGC through the LAN switch. In this case, Q.921 is not used between the AG and MGC; instead, Q.931 and IUA are used.
- The MGC converts the IUA packets containing Q.931 signaling packed by using IUA to Q.931 signaling. Also, the MGC sends Q.931 signaling to the peer end through SCTP links. This is the entire process of ISDN call signaling.

For the SIP protocol:

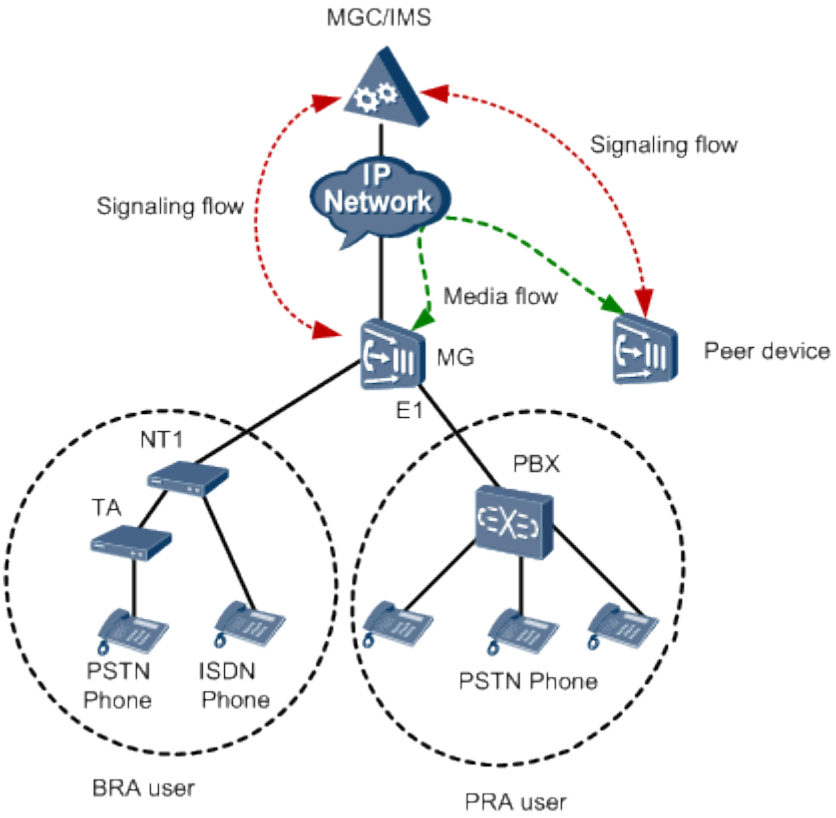
- Data transmission from ISDN user terminals to voice boards uses Q.921 and Q.931.
- The AG terminates Q.921 and Q.931 message, converts them into SIP messages and sends SIP messages to the IMS. Then the IMS sends SIP messages to the peer end. This is the entire process of ISDN call signaling.

23.9.3 Call Flow of ISDN

ISDN System Structure

Figure 23-67 shows the ISDN system structure.

Figure 23-67 ISDN System Structure



The ISDN users include the BRA users and PRA users.

- The BRA users can connect the ISDN telephone with the NT1 directly, or connect the common telephone through the TA. On the MG side, the BRA users access to the network through the BRA port. Connect the NT1 and MG with the ordinary telephone line.
- The PRA users access the network through the E1 port with the PBX. Connect the PBX and the gateway with the E1 cable.

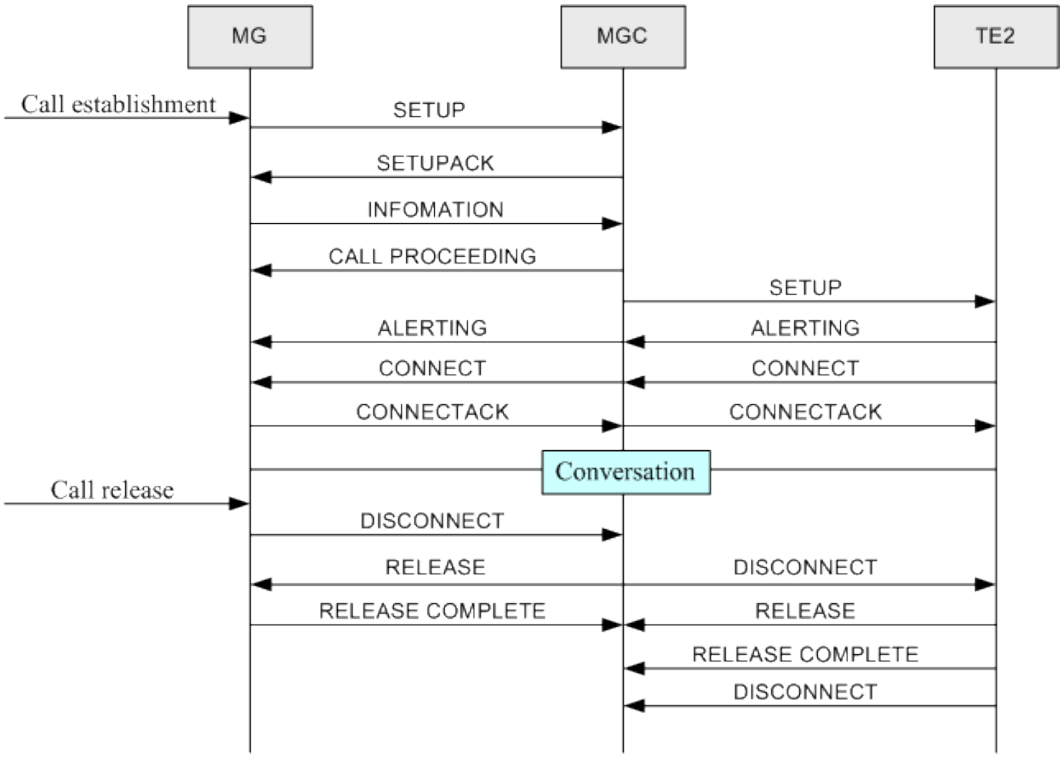
ISDN Call Control Process

- For the H.248 protocol or the MGCP protocol, the ISDN uses primitive Q931 for call control. Between the gateway and the NT1, a Layer 2 link that complies with the Q921 protocol is set up between the gateway and the PBX to carry the Q931 message. The IUA link is set up between the gateway and the softswitch to carry the Q931 message.
- For the SIP protocol, the gateway converts the ISDN message to the SIP message and then the ISDN uses the SIP protocol for call control. The gateway between the ISDN terminal and the IMS implements signaling conversion and service control. Specifically,

the gateway translates the DSS1 signaling to the SIP signaling and then sends the signaling to the IMS. Generally, the IMS transparently transmits the SIP signaling to the peer ISDN user. Then, the peer ISDN user translates the SIP signaling to the DSS1 signaling and then sends the signaling to the ISDN terminal. This process enables two ISDN terminals to be connected with each and therefore the ISDN service is provisioned.

Figure 23-68 shows the ISDN call control process.

Figure 23-68 ISDN call control process



- For the H.248 protocol or the MGCP protocol, the gateway does not process the primitive Q931 but takes out the primitive terminal Q931 from the Q921 message, encapsulates the Q.931 to the IUA message, and then sends to the softswitch. Resources are not assigned to the primitive Q931.
- For the SIP protocol, the gateway converts the ISDN message to the SIP message.

The call process includes two sections: call setup and call disconnection.

- The call setup process is as follows:
  a. The host hooks off and initiates a call setup.
  b. The softswitch responds "SETUP_ACK", and applies more call information, such as the called number.
  c. The calling party dials, and the number is carried by the primitive IMFORMATION to the softswitch.
  d. The softswitch responds "CALL PROCEEDING", and the call is setting up.
  e. The softswitch applies sending setup to the called party to set up a call.
  f. After receiving the call, the called party starts ringing and sends "ALERTING". If the "ALERTING" reaches the calling party, the call is connected.

g. The called party hooks off and sends "CONNECT". If the "CONNECT" reaches, the call is connected.

h. The calling party responds "CONNECT_ACK". The call setup is complete.

- The call disconnection process is as follows:

a. One party hooks on, and sends "DISCONNECT".

b. The softswitch sends "DISCONNECT" to the other party, and sends "RELEASE" to the party who hooks on.

c. The party who hooks on finishes the call disconnection, and sends "RELEASE_COMPLETE" to the softswitch.

d. After receiving the disconnection, the other party sends "RELEASE" to the softswitch.

e. The softswitch responds "RELEASE_COMPLETE".

f. The other party hooks on, and sends "DISCONNECT". The call disconnection is complete.

23.9.4 The Principles of ISDN BRA

Figure 23-69 and Figure 23-70 show the principles of the ISDN BRA.

Figure 23-69 H.248 protocol principles of the ISDN BRA

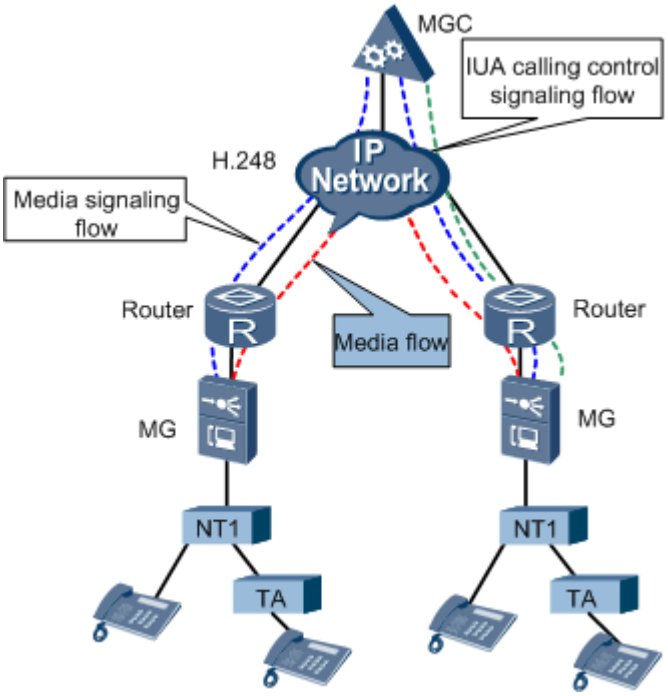
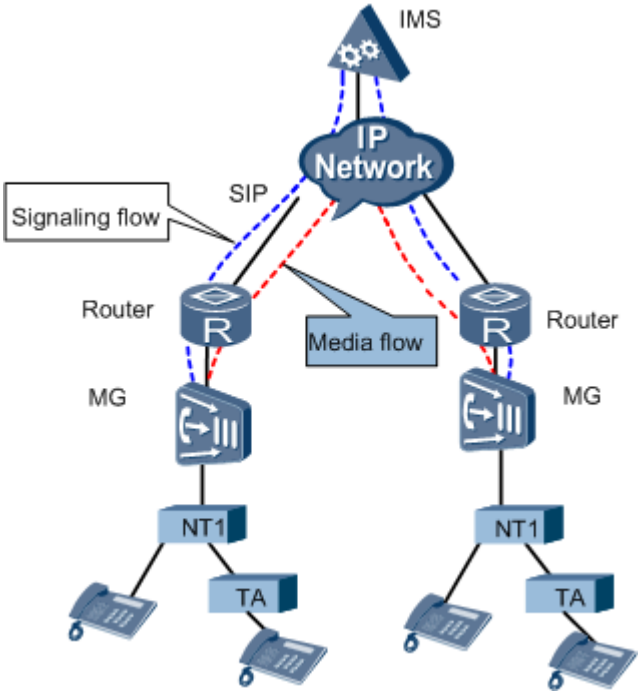


Figure 23-70 SIP protocol principles of the ISDN BRA



User Access

Entering the AN from the MG side, the BRA user call from the deactivated state experiences four stages: activation, TEI application, Layer 2 link setup, and Layer 3 call control. If the port terminal is activated, the TEI is distributed, or the Layer 2 link is set up, skip to next stage.

Call Control

- For the H.248 protocol or the MGCP protocol:
  - According to the signaling round-trip control, the call signaling on the MG is sent to the softswitch through the IUA (as the green line in the figure). The softswitch delivers the media control information through the H.248 protocol/MGCP protocol, and controls the resources on the MG (as the blue line in the figure), such as the B channel, context (H.248), and terminal.
  - Create an IUA service environment on the MG and MGC sides. Bear the Q.931 signaling on the DSL board to the SCTP link, pack the signaling through the IUA protocol stack, and then send the packet to the MGC. Switch the Q.931 signaling on the MGC side. The MGC sends the Q.931 signaling to the peer end through the SCTP link to perform ISDN signaling call.
- For the SIP protocol:
  - The ISDN service is provisioned in the IMS network by mapping and interaction between the SIP signaling and the DSS1 signaling.
  - The gateway converts the ISDN message to the SIP message and then the ISDN uses the SIP protocol for call control. The gateway between the ISDN terminal and the IMS implements signaling conversion and service control. Specifically, the gateway translates the DSS1 signaling to the SIP signaling and then sends the signaling to the IMS. Generally, the IMS transparently transmits the SIP signaling

to the peer ISDN user. Then, the peer ISDN user translates the SIP signaling to the DSS1 signaling and then sends the signaling to the ISDN terminal. This process enables two ISDN terminals to be connected with each and therefore the ISDN service is provisioned.

Working Mode

The BRA working modes include point to multipoint (P2MP) and point to point (P2P).

- In the P2MP mode, one NT1 can connect to multiple terminals. Multiple Layer 2 links can be created at the same time, and up to two users can call simultaneously. If no call service exists, the system can be deactivated automatically to save the power.
- In the P2P mode, one NT1 can connect to one terminal only. The Layer 2 link is always set up to ensure the service bearing at any moment. No matter whether the call service exists, the link is activated.

Terminal Power Supply Mode

The BRA power supply is to provide power for the terminal. Two terminal power supply modes are provided:

- Local power supply: The terminal applies battery or connects to the power supply.
- NT1 power supply: The terminal uses the NT1 power supply only. The NT1 power supply falls into two categories:
  - Local power supply: The NT1 connects to the local power supply.
  - Gateway power supply: Configure the remote power supply attribute of the BRA port on the gateway.

Terminal Identifier Distribution

In the P2MP mode, if the physical line of the BRA user is activated, one BRA port can connect multiple terminals. A terminal equipment identifier (TEI) is needed to identify the terminal.

The TEI can be specified by the terminal, or distributed on the network side.

- The TEI that the terminal specifies ranges 0-63.
- The TEI on the network side is distributed by the subscriber board, ranging 64-126.
- The 127, as a multicast TEI, is used when the BRA user is called (all the users under the same port share the same telephone number). When the destination terminal is unknown, the connections to all the terminals are initiated.
- In the P2P mode, the terminal TEI is 0.

23.9.5 The Principles of ISDN PRA

The PRA call process is the same as the BRA call process. For the BRA call process, refer to 23.9.4 The Principles of ISDN BRA.

- One PRA user has 32 time slots with the rate of 64 kbit/s. Among the 32 time slots, time slots 1-15, 17-31 are for the B channel, time slot 16 is for the D channel, and time slot 0 is for frame synchronization.
- For a PRA user, the TEI of the L2 link is 0.

- For a PRA user, the working mode and power supply mode are not included. The terminal is powered by the PBX.

23.9.6 ISDN Standards and Protocols Compliance

This topic provides the reference documents of the ISDN:

- ITU-T Q.920 ISDN user-network interface data link layer General aspects
- ITU-T Q.921 ISDN user-network interface - Data link layer specification
- ITU-T Q.930 Digital Subscriber Signalling System No.1 (DSS 1) -ISDN User-Network Interface Layer 3 - General Aspects
- ITU-T Q.931 ISDN user-network interface layer 3specification for basic call control
- ITU-T H.248 Media gateway overload control package
- RFC3435 Media Gateway Control Protocol (MGCP) Version 1_0
- RFC3660 Basic Media Gateway Control Protocol (MGCP) Packages
- RFC3661 Media Gateway Control Protocol (MGCP) Return Code Usage
- ITU-T G.961 Digital transmission system on metallic local lines for ISDN basic rate access

23.10 R2 Access

R2 access enables the MA5600T/MA5603T/MA5608T to be interconnected with a private branch exchange (PBX) through the R2 signaling and helps to provide access services for users over the common twisted pairs. As a type of channel associated signaling (CAS), R2 signaling is the international standard signaling based on E1 digital networks.

23.10.1 Introduction to the R2 Feature

Definition

R2 signaling is a type of channel associated signaling (CAS) and it is also the international standard signaling based on E1 digital networks. Timeslot 16 in the R2 signaling is reserved for transmitting the signaling of the voice channel.

Actually, there is no agreed standard for R2 signaling. ITU-T Recommendation Q.400-Q.490 define the R2 signaling standards but different countries and regions have developed their own standards.

Purpose

The MA5600T/MA5603T/MA5608T connects an R2 PBX to the next generation network, achieving transformation of the public switched telephone network (PSTN) to the NGN.

23.10.2 R2 Principles

R2 signaling is inter-office channel associated signaling, and applies to international/national networks. R2 signaling is specified on both analog and digital transmission systems. R2 signaling contains line signaling and register signaling. Line signaling is available in three forms: DC line signaling, inband single-frequency pulse line signaling, and digital line signaling. For multi-end route transmission, the link-by-link forwarding mode is used.

Register signaling can be transmitted in multi-frequency compelled (MFC) mode and dual tone multiple frequency (DTMF) mode.

Line Signaling

Line signaling is primarily used for monitoring the occupation, release and congestion states of a trunk line. Line signaling is classified into analog line signaling and digital line signaling. The MA5600T/MA5603T/MA5608T supports only the digital line signaling which will be described in detail in the following paragraph.

The digital line signaling uses timeslot 16 of the PCM for transmitting line signaling at a rate of 2048 kbit/s. To transmit line signaling of 30 voice channels, 16 frames form a multiframe. Timeslot 16 of frame 0 in the multiframe is used for multiframe synchronization. The first four bits of timeslot 16 in frame 1 correspond to the first voice channel, while the last four bits correspond to the 16th voice channel and so on.

Register Signaling

Register signaling is the signaling transmitted over a voice channel after the line signaling occupies the voice channel. Register signaling, including the selection signaling and service signaling, is used to transmit the control signals for a voice channel connection, such as managing the telephone network and selecting the route and the called party. The MFC register signaling will be described in detail.

The MFC register signaling includes the forward signaling and backward signaling, both of which are consecutive. The forward signaling is used to transmit the address information and control the indication information, while the backward signaling is used for acknowledgement and control. When transmitting a number, the transmit end stops transmitting the forward signaling only after receiving the acknowledgement from the backward signaling. Similarly, the receive end stops transmitting the backward signaling only after confirming that the transmission of the forwarding signaling is stopped. A control period can be divided into four steps, which is listed as follows:

- Step 1: The user side sends the forward signaling.
- Step 2: The network side receives the forward signaling and returns the backward signaling.
- Step 3: The user side receives the backward signaling and stops sending forward signaling.
- Step 4: The network side finds that the transmission of forward signaling is stopped, and stops sending the backward signaling.

The MFC register signaling uses the arithmetic frequency with 120 Hz as the common difference. This section takes the definition in Q.441 as an example, the forward signaling uses the high-frequency group (1380 Hz to 1980 Hz) and selects two out of six frequencies (1380 Hz, 1500 Hz, 1620 Hz, 1740 Hz, 1860 Hz, and 1980 Hz) for encoding. Up to 15 signaling combinations can be formed. The backward signaling uses the low-frequency group (780 Hz to 1140 Hz) and selects two out of four frequencies (780 Hz, 900 Hz, 1020 Hz, and 1140 Hz) for encoding. Up to six signaling combinations can be formed. To expand the signaling capability, the forward signaling is divided into group I forward signaling and group II forward signaling, while the backward signaling is divided into group A backward signaling and group B backward signaling.

- Register signaling must always start with group I forward signaling. Group I forward signaling includes the information of the country code, echo suppressor indicator (I-11) and address signal (number: 1-9).

- Group II forward signaling is the calling party's category signaling sent by an outgoing R2 register. It is used to reply to backward signal A-3 (the address-complete signal) or A-5 (the request signal for a calling party's category), and send the national or international calling information.
- Group A backward signaling is used to acknowledge the group I forward signaling and under some conditions, group II forward signaling, such as acknowledging the calling party's type and group II forward signals.
- Any group B backward signaling acknowledges the group II forward signaling and is always preceded by the address-complete signal A-3. Signal A-3 indicates that the incoming R2 register has received all the required forward signals from the outgoing R2 register.

Register signaling is the in-band signaling (the frequency is within the voice frequency band). Therefore, the register signaling is transmitted over the voice channel.

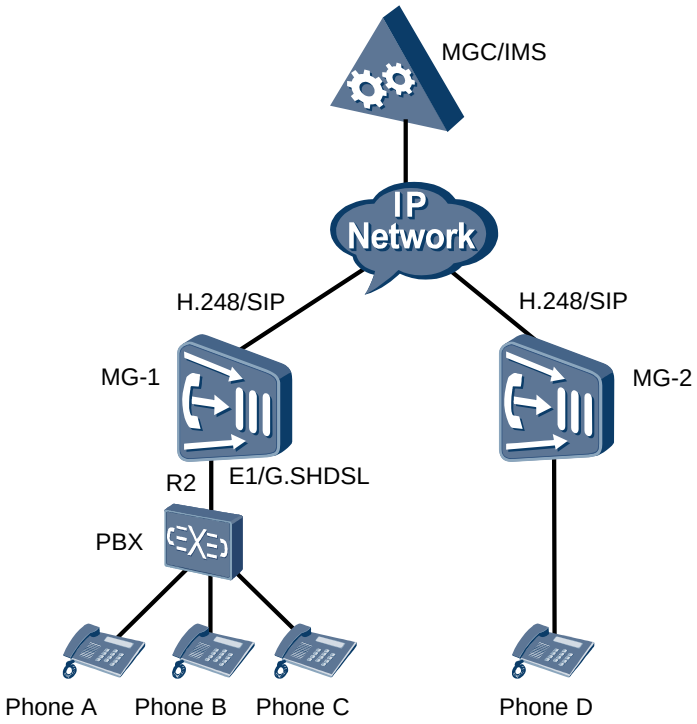
R2 Application in Access Networks

This section takes H.248 protocol as an example. A PBX is connected to the next generation network (NGN) through R2 signaling. Interfaces between an MG and an MGC use the H.248 protocol, and interfaces between an MG and a PBX use the R2 protocol. Figure 23-71 shows the networking of connecting the R2 PBX to the NGN network.

To connect a PBX to the NGN network through R2 signaling, MGs need to perform the conversion between R2 signaling and H.248 signaling.

- In the upstream direction, the MA5600T/MA5603T/MA5608T terminates R2 signaling transmitted from the PBX, converts R2 signaling to H.248 signaling, and sends the converted signaling to the softswitch.
- In the downstream direction, the MA5600T/MA5603T/MA5608T terminates H.248 signaling transmitted from the softswitch, converts H.248 signaling to R2 signaling, and sends the converted signaling to the PBX.

Figure 23-71 Network example of connecting the R2 PBX to the NGN network



23.10.3 R2 Standards and Protocols Compliance

The reference standards and protocols of this feature are as follows:

- draft-manyfolks-megaco-caspackage-01.txt
- draft-laha-megaco-cas-mntc-00.txt
- draft-ietf-megaco-r2package-03.txt
- ITU-T H.248.25 Gateway control protocol: Basic CAS packages
- Specifications of Signaling System R2, Q.400 to Q.490, Blue Book, CCITT

23.11 FoIP

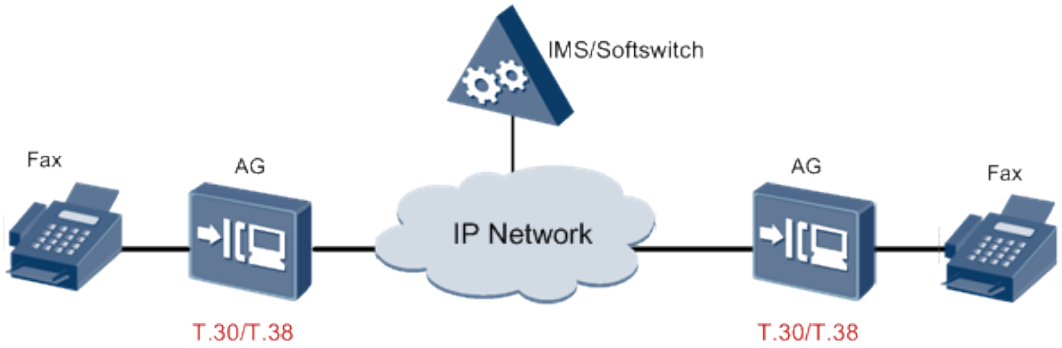
Fax over Internet Protocol (FoIP) is a fax service provided on an IP network or between an IP network and a traditional PSTN. Fax service is a data service that is widely applied on the PSTN network.

23.11.1 What Is FoIP

Definition

Fax over Internet Protocol (FoIP) is a fax service provided on an IP network or between an IP network and a traditional PSTN. Figure 23-72 shows the network application of FoIP.

Figure 23-72 FoIP network application



Basic Process

The basic process of fax can be described as follows:

1. The fax transmitting machine scans a page to obtain image information.
2. The fax transmitting machine digitalizes and compresses the image signals.
3. The fax transmitting machine modulates the image signals into analog signals, and transmits the signals to the fax receiving machine through common subscriber lines (as defined in the T.30 protocol).

Commonly Used Protocols for Fax

The following protocols are commonly used for fax:

- T.30: It is a fax protocol based on the PSTN network. T.30 defines in detail the process for transmitting fax signals on a PSTN network. It also defines the modulation mode (V.17/V.21/V.27/V.29/V.34) and transmission format (HDLC) of data, and the physical standard for fax signals. The T.30 fax messages and data can be transmitted transparently between AGs. This is called the T.30 transparent transmission mode. The quality of fax in this mode may not be high due to packet loss, delay, and packet disorder on the IP network.
- T.38: It is a protocol that defines the process for carrying fax services over a packet IP network. Serving as a supplementary protocol to T.30, T.38 implements packet encapsulation based on T.30 in order to adapt to IP applications. In T.38 fax service, the T.38 redundancy mechanism is used to ensure service quality.

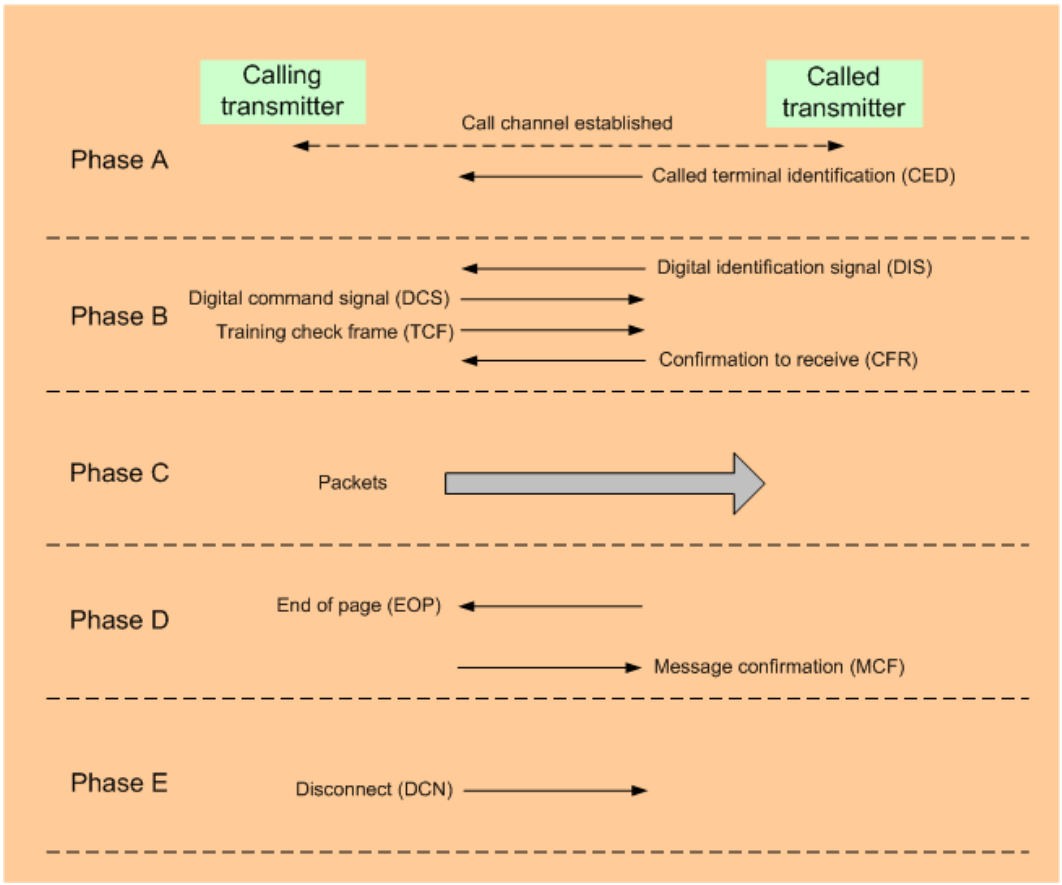
Typical Fax Process Defined in T.30

T.30 defines 5 phases for a typical fax process, as shown in Figure 23-73.

- Phase A: A call for fax is established. This phase is similar to a telephone call establishment phase.
- Phase B: The earlier phase of the packet transmission process. In this phase, the devices at both ends perform capability negotiation and training.
- Phase C: This is the packet transmission phase and also the packet transmission control phase.
- Phase D: The later phase of the packet transmission process. In this phase, packets are verified, errors are corrected, and multiple pages are continuously transmitted.

- Phase E: The call for fax is released.

Figure 23-73 Typical fax process defined in T.30



23.11.2 Classification of FoIP

Fax over IP (FoIP) is a faxing service provided over an IP network or between an IP network and a traditional PSTN network. The fax machine can be regarded as a special modem. In the FoIP negotiation, the modem negotiation is performed before the fax negotiation. FoIP services can be classified based on the transmission real-time performance, or based on the transmission mode.

Classification

Classification Based on Transmission Real-time Performance

Based on the real-time performance, FoIP can be classified in store-and-forward FoIP and real-time FoIP. The difference between the two modes lies in whether communication between the gateway and the fax machine is real-time on the IP side. On the PSTN side, communication of the two FoIP modes is real-time.

- Store-and-forward FoIP: In this mode, fax information is stored and then forwarded to the IP network, as defined in the T.38 protocol.
- Real-time FoIP: In this mode, communication during the entire fax process is carried out in real time, as defined in the T.38 protocol.

NOTE

Huawei access gateway (AG) supports real-time FoIP.

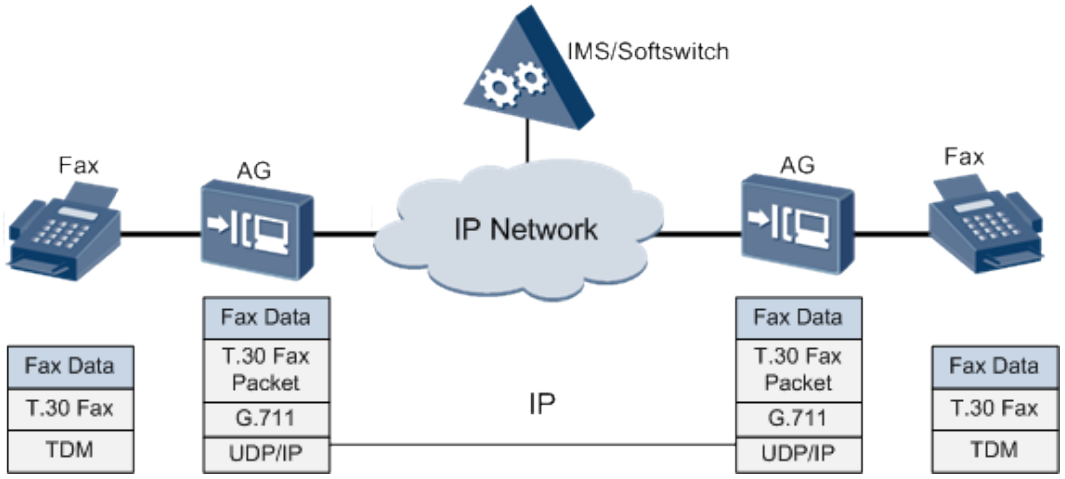
Classification Based on Transmission Mode

According to the transmission protocol used, there are two modes of fax services carried over the IP network: the T.30 transparent transmission mode and the T.38 transmission mode.

T.30 transparent transmission mode

In this mode, T.30-defined fax messages and data are transparently transmitted in an AG or between AGs. Figure 23-74 shows the T.30 transparent transmission mode.

Figure 23-74 T.30 transparent transmission mode



The advantages and disadvantages of the T.30 transparent transmission mode are as follows:

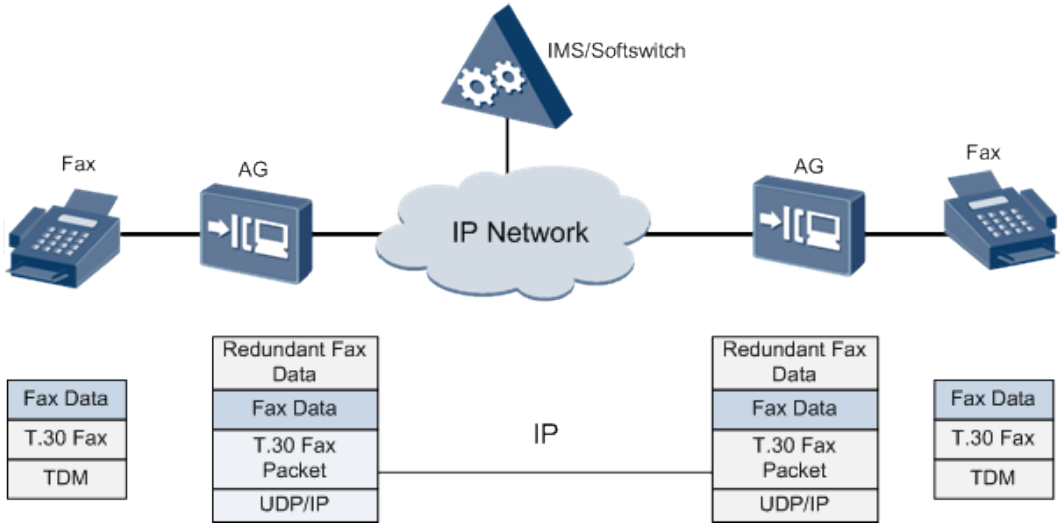
- Advantages: consumes less DSP resources, and is less dependent on the softswitch.
- Disadvantages: weak resistance against interference from the network, and does not provide reliable guarantee for fax quality. Stability of the T.30 transparent transmission mode can be improved using the RFC 2198 and 10-ms packetization technologies. For more details, see 23.14.8 Fax/Modem Quality Enhancement.

T.38 transmission mode

The T.38 transmission mode is shown in Figure 23-75. T.38 fax supports two rate negotiation modes: end-to-end negotiation and local negotiation. The difference between the two negotiation modes lies in whether the rate training signals need to be transmitted from the transmitting AG to the receiving AG.

- When the rate training signals need to be transmitted from the transmitting AG to the receiving AG, it is end-to-end negotiation.
- When the rate training signals are terminated and generated by the transmitting AG, it is local negotiation. If the local negotiation mode is used, the maximum rate supported by the AG should be considered. That maximum rate is reflected by the maximum fax rate supported by the digital signal processor (DSP).

Figure 23-75 T.38 transmission mode



The advantages and disadvantages of the T.38 transmission mode are as follows:

- Advantages: provides a redundancy mechanism for transmitting data packets, and does not have strict requirements on network quality (able to process the fax service even with a 20% packet loss rate on the network).
- Disadvantages: The DSP chip of the AG needs to participate in parsing the T.38 signals. Because there are various types of terminals on the network, interoperability problems may occur.

Differences Between High-speed Fax and Low-speed Fax

The main differences between high-speed fax and low-speed fax include the following:

- Standards applied. High-speed fax applies the V.8 data transmission process, while low-speed fax applies a fax process defined by the T.30 protocol. In addition, some low-speed fax terminals may use earlier standards.
- Range of rates supported. High-speed fax supports a rate range of 2400 bit/s-33600 bit/s, while low-speed fax supports a rate range of 2400 bit/s-14400 bit/s.
- Upstream transmission modes used. High-speed fax can use only the T.30 transparent transmission mode. In other words, for an AG, only the high-speed modem T.30 transparent transmission mode can be used for fax services. Low-speed fax can use the T.30 transparent transmission mode or T.38 transmission mode, according to data configuration of the fax terminal.
- Error correction requirements. For high-speed fax, the error correction function is a mandatory requirement; for low-speed fax, the error correction function is optional.
- Echo cancellation (EC) requirements. High-speed fax requires the EC function to be disabled because a high-speed fax terminal already has an inherent EC mechanism; low-speed fax requires the EC function to be enabled.

23.12 MoIP

Modem over Internet Protocol (MoIP) is a technology for providing modem services over an IP network or between an IP network and a PSTN network.

23.12.1 What Is MoIP

Definition

The term "modem" is abbreviated from modulator and demodulator. Modem service is a data service that is widely applied on the PSTN network. In its earlier application, modem service is mainly used for point-to-point dialing and Internet dialup. Later, the service scope is expanded to cover point of service (POS) machine connection, alarming, and lottery machine connection.

Modem over Internet Protocol (MoIP) is a technology for providing modem services over an IP network or between an IP network and a PSTN network.

The dialup mode of modems can be pulse dialup and tone dialup.

- In the traditional modem dialup scenario, PSTN users directly dial numbers, and the media stream passes through only the narrowband channel.
- MoIP differs from the traditional modem dialup service in that MoIP users are next generation network (NGN) access gateway (AG) users, and that modem negotiation and media stream transmission between the AG and the trunk gateway are performed based on IP.

MoIP Services

Modem services are usually transmitted using the V.32/V.34/V.90/V.92 protocol. Different modem protocols support different ranges of rates. In practical service application, the transmission rate is usually the result negotiated by the two modems at both ends.

- V.32 supports a maximum rate of 14.4 kbit/s.
- V.34 supports a maximum rate of 33.6 kbit/s. V.34 is the modem protocol most commonly used. It defines a negotiation process that is the same as the high-speed fax process. One application example of V.34 is the POS machine service.
- V.90/V.92 supports a maximum rate of 56 kbit/s and is usually used for Internet dialup services.

Modem services are classified into high-speed modem services and low-speed modem services. The negotiation process of high-speed modem services is the same as that of high-speed fax services.

- Based on the service rate, services with rates lower than or equal to 14.4 kbit/s are low-speed modem services, and services with rates higher than 14.4 kbit/s are high-speed modem services.
- Based on the modem's requirements on the network, modems that require the network device involved to disable the echo cancelation function are high-speed modems, and modems that do not require so are low-speed modems.

23.12.2 Principle of MoIP

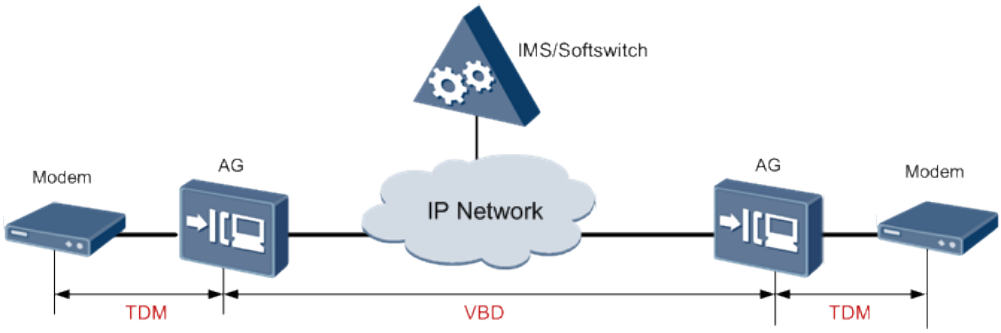
Just like VoIP, MoIP supports the AG-based PSTN-IP-PSTN network structure, and also supports the PSTN-IP network structure.

MoIP Transmission Modes

MoIP has two transmission modes:

- Transparent transmission mode, which is also called the voice-band data (VBD) transparent transmission mode. In this mode, the AG uses the G.711 mode to encode and decode the modem signals, and processes the signals like processing common Real-time Transfer Protocol (RTP) data. In other words, the AG does not process the modem modulation signals, which are transparently transmitted on the IP network through VoIP channels. Figure 23-76 shows the MoIP transparent transmission mode.

  Figure 23-76 MoIP transparent transmission mode

  - High-speed modem: with voice activity detector (VAD) and echo canceller (EC) disabled.
  - Low-speed modem: with VAD disabled and EC enabled.
- Redundancy mode, which is also called the relay mode.

NOTE

Currently, Huawei AGs support only the transparent transmission mode for MoIP.

Enhanced MoIP

The MoIP service carried in transparent transmission mode has high requirements on the bearer network. Jitter, delay, and packet loss on the network will affect modem services significantly. In case of poor network quality, the MoIP service can be enhanced using the RFC 2198 and 10-ms packetization technologies. Using these technologies, the connection success ratio can be improved for modems, and the modem disconnection ratio will be reduced as well. For more details, see 23.14.8 Fax/Modem Quality Enhancement.

23.13 IP Z Interface Extension

IP Z interface extension is that the analog interface between an accsee device and a PBX extends to the remote place through the IP network.

23.13.1 Introduction to IP Z Interface Extension

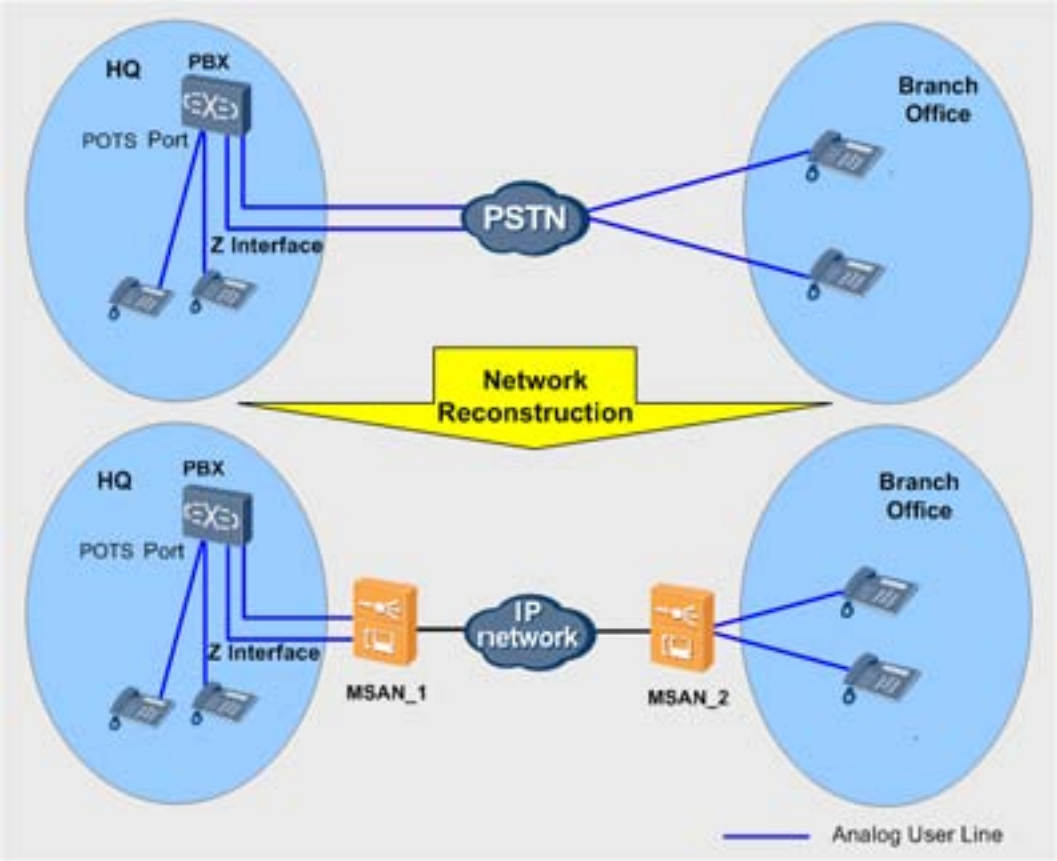
Context

The PSTN network can be used to provision the Z interface extension private line service for the headquarters (HQ) and branch offices of businesses. The service enables users at the branch offices to have the same calling experience as users at the HQ through phones. As devices on existing PSTN networks are approaching the end of their life cycles, many carriers are reconstructing the PSTN networks.

Figure 23-77 illustrates the Z interface extension service network before and after the PSTN network reconstruction. After the reconstruction, the Z interface extension service that used to be carried by the PSTN network is carried by an IP network.

The access devices, one located at the HQ and the other at the branch office, are interconnected through the IP network. The voice signals on POTS lines are carried through RTP streams when transmitted over the IP network. The call signaling on POTS lines, such as offhook and onhook signaling, is carried through RFC2833 packets.

Figure 23-77 Z interface extension service network before and after PSTN reconstruction



Definition

The Z interface refers to the analog interface between an access device and an exchange.

Z interface extension means the extension of POTS user signals. Analog POTS user signals are converted into digital signals at the Tx end, transmitted over the network, and restored into analog signals at the Rx end. In this feature, the extension of the Z interface is carried over an IP network, so the feature is called IP Z interface extension.

Purpose

IP Z interface extension is intended for the following purposes:

- Connect the private line users at the branch offices of an enterprise that has a small volume of analog phone service requirements to the enterprise's local exchange.
- Extend the calls of analog phones under a PBX to a remote location, thereby saving long-distance call fee.
- Connect users at a remote location to a local exchange.

Hardware support

Boards on the HQ-side MSAN (such as MSAN_1 in Figure 23-77) that provide the IP Z interface extension service are the FXO board. Boards on the branch office-side MSAN (such as MSAN_2 in Figure 23-77) for the same purpose are the FXS board. Table 23-15 provides the details.

Table 23-15 Hardware support for the IP Z interface extension feature

Product	Board Type	Board Name
MA5600T	Control board	SCUB SCUN SCUK SCUH
	Backplane	H802MABC
	FXO board	H80AATRB
	FXS board	H801ASPB H808ASPB H809ASPB H80BCAME H806CAME H80BCVME H806CCPE
MA5603T	Control board	SCUB SCUN SCUK SCUH
	Backplane	H802MABO
	FXO board	H80AATRB

Product	Board Type	Board Name
	FXS board	H801ASPB H808ASPB H809ASPB H80BCAME H806CAME H80BCVME H806CCPE
MA5608T	Control board	H801MCUD H801MCUD1
	Backplane	H801MABR
	FXO board	H80AATRB
	FXS board	H801ASPB H808ASPB H809ASPB H80BCAME H806CAME H80BCVME H806CCPE

Limitations

IP Z interface extension is a technology proprietarily owned by Huawei. Therefore, the MSANs on the FXO side and the FXS side must be Huawei MSANs.

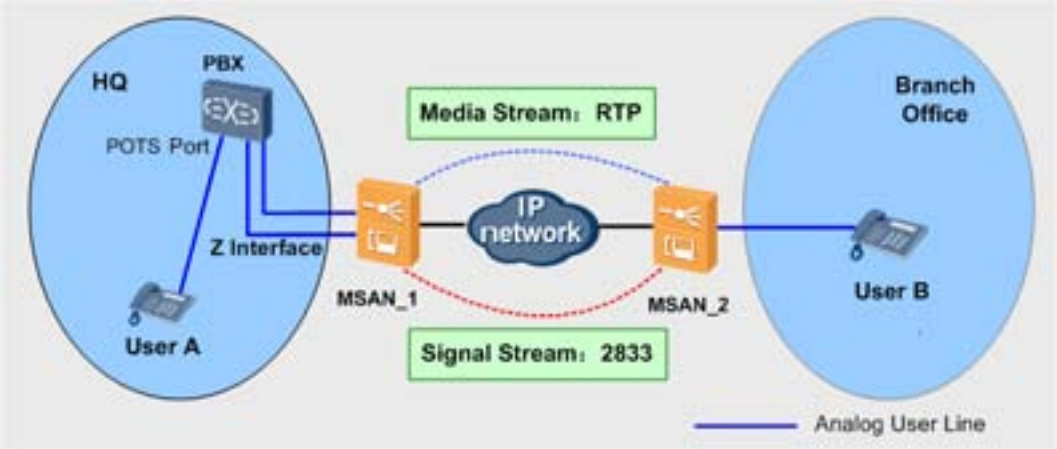
23.13.2 Principle of IP Z Interface Extension

The typical IP Z interface extension service includes a Z interface, the IP transmission network, and a POTS port.

Figure 23-78 demonstrates the implementation principle of IP Z interface extension. In this network diagram, two access devices, MSANs, one located at the headquarters (HQ) and the other at the branch office, are interconnected through the IP network. Such a networking model eliminates the limitations on the line type and line transmission distance, and covers a wider range of users. MSAN_1 at the HQ converts the analog signals of the local exchange (PBX) into digital signals using the analog-to-digital conversion mechanism. The voice signals on POTS lines are carried through RTP streams when transmitted over the IP network. The call signaling on POTS lines, such as signaling for offhook, onhook, hookflash pressing, and ringing, is carried through RFC2833 packets. In this way, the voice signals and call signaling are transmitted using the IP network to the POTS port of the remote access device MSAN_2. MSAN_2 restores the analog signals from the digital signals using the digital-to-analog mechanism, thus extending POTS signals from the Z interface of MSAN_1 to the POTS port of MSAN_2.

The IP Z interface extension services enables user A and user B to have the same user experience when making outgoing and incoming calls. The Z interface of MSAN_1 and the POTS port of MSAN_2 are in one-to-one mapping and are configured on a 1:1 basis.

Figure 23-78 Principle of IP Z interface extension



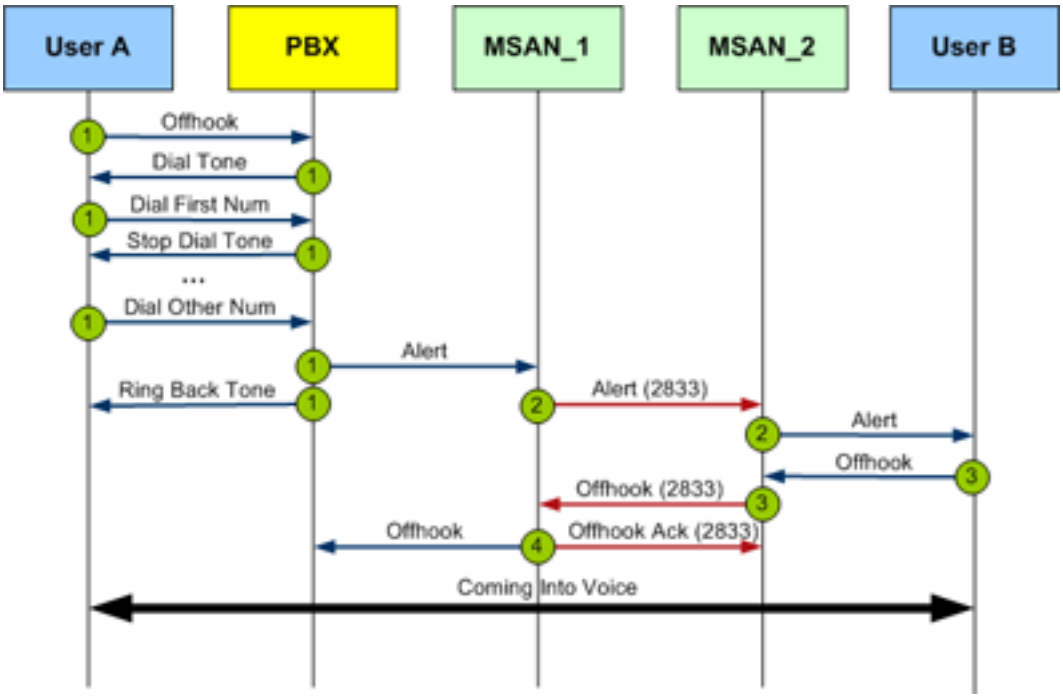
23.13.3 Call Service Flows of IP Z Interface Extension

IP Z interface extension supports many service flows.The following describes the call service flows which include outgoing call service flow for POTS user A, outgoing call service flow for user B with IP Z interface extension, call release flow for POTS user A and call release flow for user B with IP Z interface extension.

Outgoing Call Service Flow for POTS User A

Figure 23-79 shows the outgoing call service flow for POTS user A.

Figure 23-79 Outgoing call service flow for POTS user A



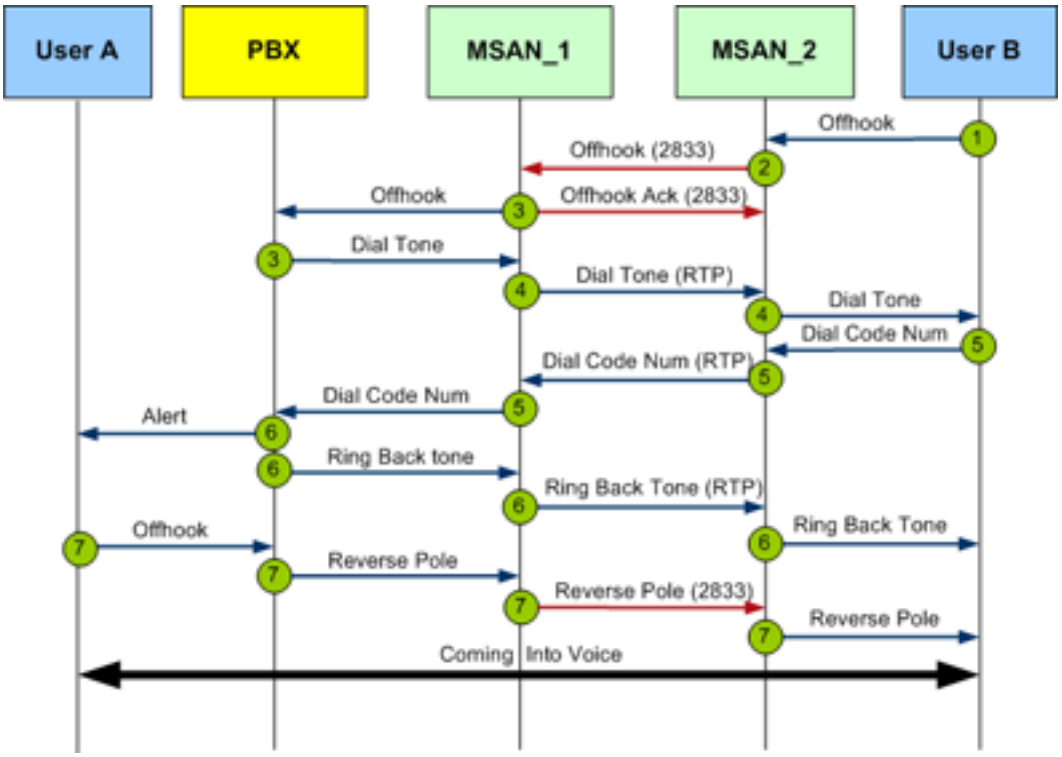
The outgoing call service flow for POTS user A is as follows:

1. User A picks up the phone, the PBX detects the offhook signal of user A and sends the dial tone to user A, and then user A starts to dial the number. After collecting the number, the PBX plays the ringing tone to the called port, and at the same time plays the ringback tone to user A.
2. After detecting the ringing tone, MSAN_1 transmits the ringing tone information to the FXS board of MSAN_2 through an RFC2833 packet. After receiving the RFC2833 packet, MSAN_2 restores the analog signals from the ringing tone signals and plays the analog ringing tones to user B.
3. After hearing the ringing tone, user B picks up the phone. After detecting the offhook signal of user B, the FXS board of MSAN_2 sends the offhook information to MSAN_1 through an RFC2833 packet.
4. After receiving the offhook information, MSAN_1 sends an offhook acknowledge message to MSAN_2 through RFC2833, and at the same time informs the PBX that user B has picked up the phone.
5. After receiving the offhook acknowledge message, the entire speech channel is set up, and the call is established between user A and user B.

Outgoing Call Service Flow for User B with IP Z Interface Extension

Figure 23-80 shows the outgoing call service flow for user B with IP Z interface extension.

Figure 23-80 Outgoing call service flow for user B with IP Z interface extension



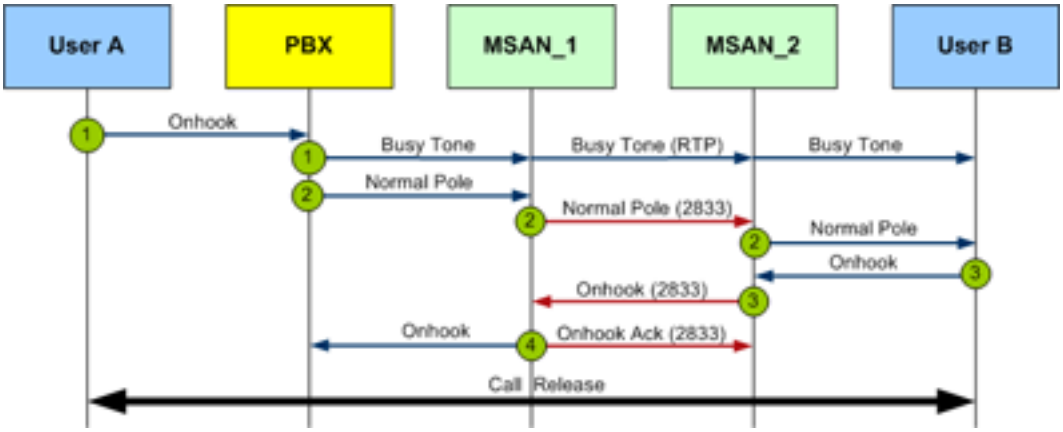
The outgoing call service flow for user B with IP Z interface extension is as follows:

1. User B picks up the phone.
2. MSAN_1 detects the offhook signal of user B and sends the offhook signal through an RFC2833 packet to inform MSAN_2 that user B has picked up the phone.
3. MSAN_2 sends an offhook acknowledgement message to MSAN_1 through RFC2833, and at the same time sends the offhook signal to the PBX. After receiving the offhook signal, the PBX plays the dial tone to MSAN_1.
4. MSAN_1 sends the received dial tone to MSAN_2 through RTP streams, and MSAN_2 sends the dial tone to user B.
5. User B dials the number. The FXS board of MSAN_2 sends the dialed number to MSAN_1 through RTP streams.
6. The PBX collects the number, plays the ringing tone to user A, and at the same time transparently transmits the ringback tone to user B through RTP streams.
7. User A hears the ringing tone and picks up the phone. The PBX detects the offhook signal of user A and sends a polarity reversal signal to MSAN_1. MSAN_1 sends the polarity reversal information to MSAN_2 through an RFC2833 packet. MSAN_2 performs polarity reversal billing on user B. By now, the entire speech channel is set up, and the call is established between user A and user B.

Call Release Flow for POTS User A

Figure 23-81 shows the call release flow for POTS user A.

Figure 23-81 Call release flow for POTS user A



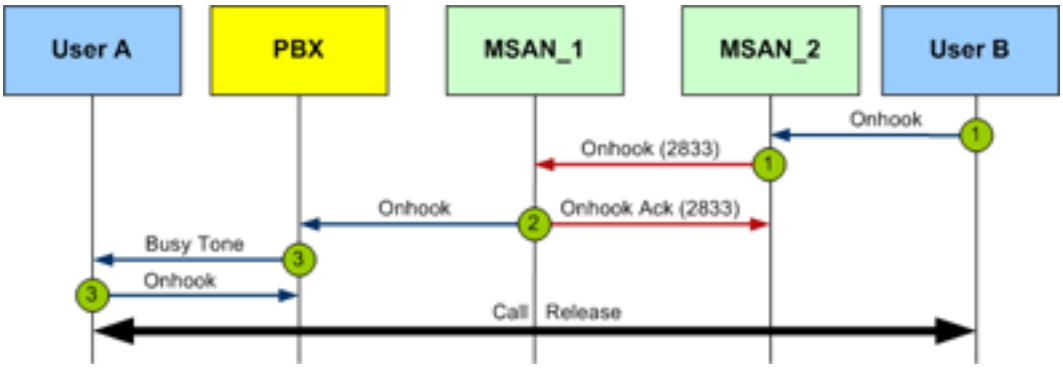
The call release flow for POTS user A is as follows:

1. User A hangs up the phone. The PBX plays the busy tone, and MSAN_1 transparently transmits the busy tone.
2. After user A hangs up the phone, the PBX restores the normal polarity. MSAN_1 detects the polarity restoration message on the port and sends the information to MSAN_2 through an RFC2833 packet. Then, MSAN_2 restores the polarity of user B.
3. User B hears the busy tone and hangs up the phone. MSAN_2 sends the onhook event to MSAN_1 through RFC2833 signaling.
4. When receiving the onhook message, MSAN_1 sends an onhook acknowledgement message to MSAN_2 through RFC2833, and at the same time sends an onhook message to the PBX.
5. After receiving the onhook message, the PBX releases the call.

Call Release Flow for User B with IP Z Interface Extension

Figure 23-82 shows the call release flow for user B with IP Z interface extension.

Figure 23-82 Call release flow for user B with IP Z interface extension



The call release flow for user B is as follows:

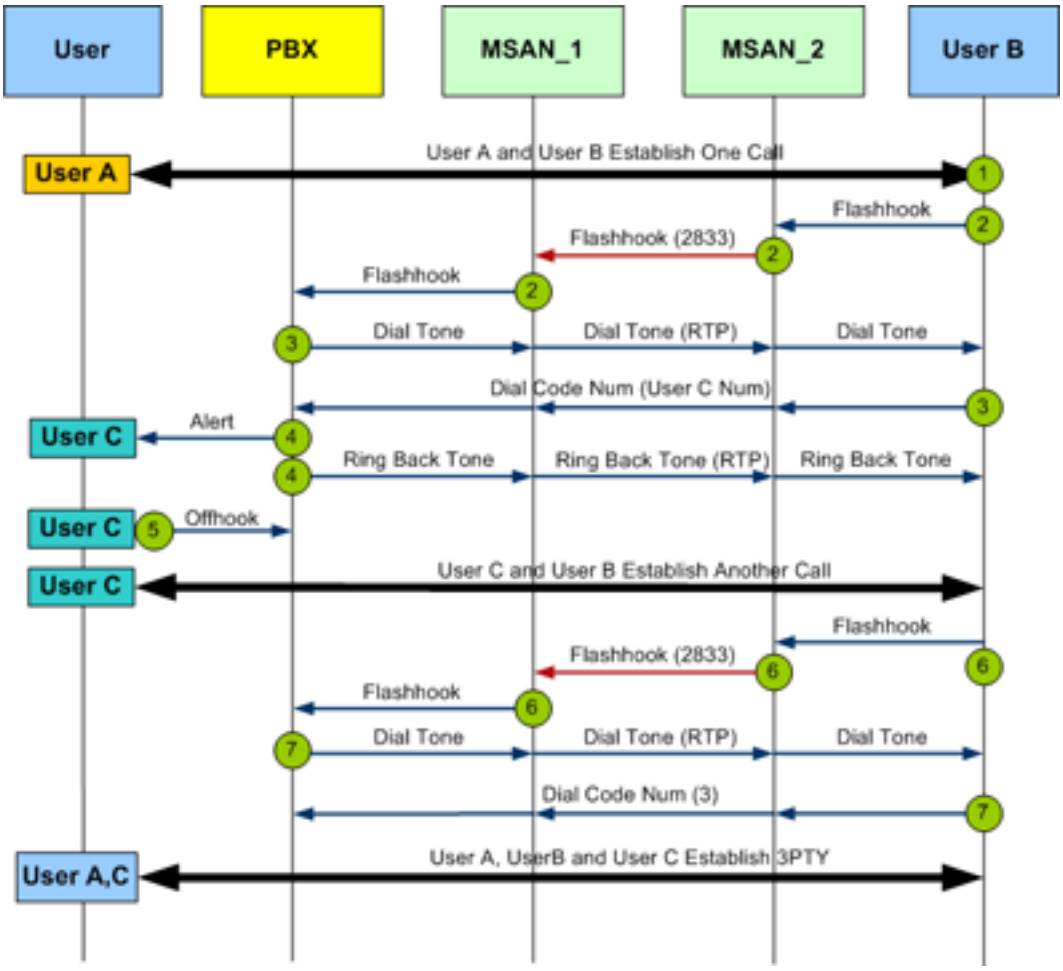
1. User B connected to the FXS board hangs up the phone. MSAN_2 detects the local onhook event and sends the event to the FXO board of MSAN_1 through RFC2833.

2. The FXO board sends an onhook acknowledgement to the FXS board through an RFC2833 packet and sends an onhook message to the PBX.
3. After receiving the onhook message, the PBX plays the busy tone to user A. User A hangs up the phone. The call is released.

Flow of Carrying New Service through IP Z Interface Extension

The IP Z interface extension feature supports news services that include call waiting and three-way calling (3WC). The following uses 3WC as an example to describe how the new service is carried using the IP Z interface extension feature. Figure 23-83 shows the service flow.

Figure 23-83 Flow of carrying the 3WC service through IP Z interface extension



The flow of carrying the 3WC service through IP Z interface extension is as follows:

1. A call has been set up between user A and user B.
2. User B needs to communicate with user C and therefore presses the hookflash. MSAN_2 detects the hookflash message, encapsulates the message as an RFC2833 packet, and sends the packet to the FXO board of MSAN_1. MSAN_1 restores the hookflash message from the RFC2833 packet and sends the message to the PBX.

3. After receiving the hookflash message, the PBX transparently transmits the dial tone to user B. After hearing the dial tone, user B starts to dial the number of user C. The dialed number is transparently transmitted to the PBX.
4. After collecting the number, the PBX plays the ringing tone to user C and plays the ringback tone to user B.
5. User C picks up the phone. The call is established between user B and user C.
6. User B presses the hookflash again. MSAN_2 sends the hookflash signal to MSAN_1 through RFC2833. MSAN_1 informs the PBX of user B's hookflash pressing.
7. After receiving the hookflash message, the PBX transparently transmits the dial tone to user B. After hearing the dial tone, user B starts to dial the DTMF number 3. The dialed number is transparently transmitted to the PBX. By now, the 3WC is established between users A, B, and C.

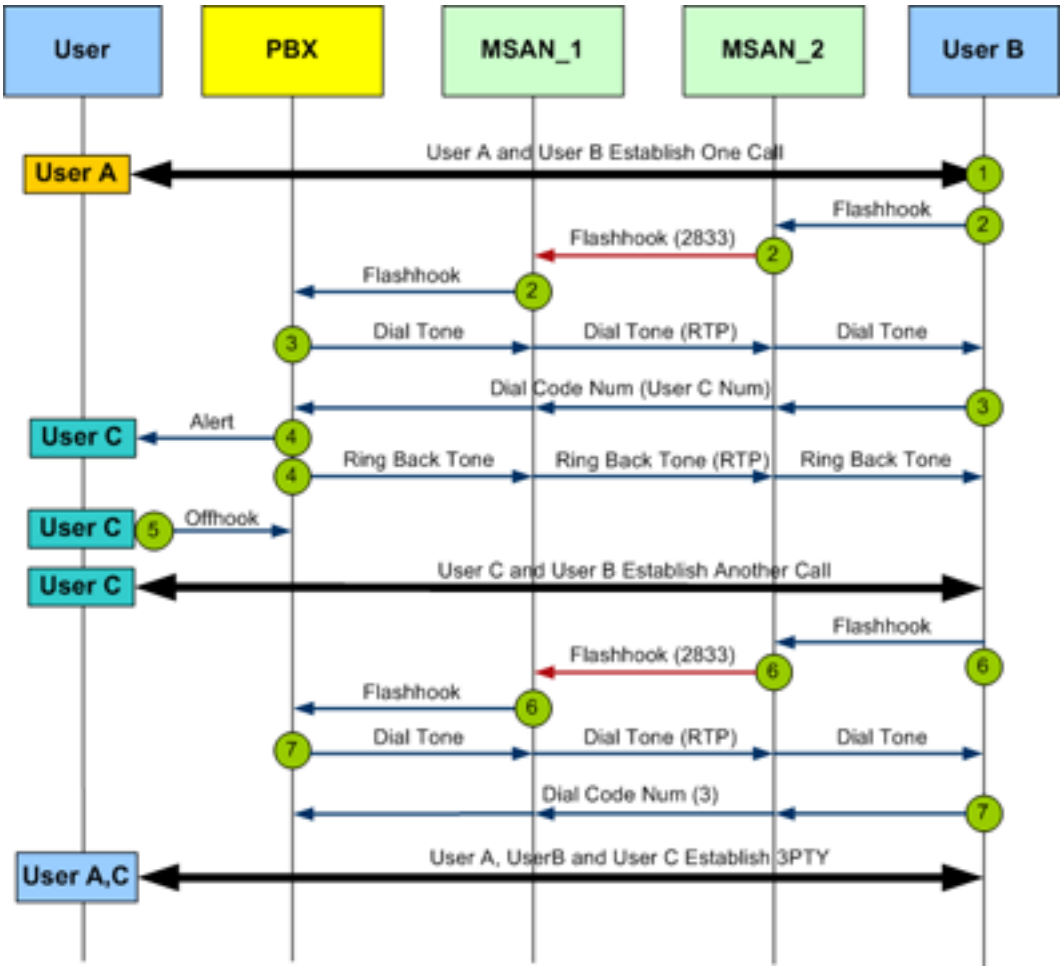
NOTE

During the time from user B pressing the hookflash for the first time to user B pressing the hookflash for the second time and dialing the DTMF number 3, user A is in the call waiting state.

23.13.4 Carrying New Service Flows of IP Z Interface Extension

The IP Z interface extension feature supports new services that include call waiting and three-way calling (3WC). The following uses 3WC as an example to describe how the new service is carried using the IP Z interface extension feature. Figure 23-84 shows the service flow.

Figure 23-84 Flow of carrying the 3WC service through IP Z interface extension



The flow of carrying the 3WC service through IP Z interface extension is as follows:

1. A call has been set up between user A and user B.
2. User B needs to communicate with user C and therefore presses the hookflash. MSAN_2 detects the hookflash message, encapsulates the message as an RFC2833 packet, and sends the packet to the FXO board of MSAN_1. MSAN_1 restores the hookflash message from the RFC2833 packet and sends the message to the PBX.
3. After receiving the hookflash message, the PBX transparently transmits the dial tone to user B. After hearing the dial tone, user B starts to dial the number of user C. The dialed number is transparently transmitted to the PBX.
4. After collecting the number, the PBX plays the ringing tone to user C and plays the ringback tone to user B.
5. User C picks up the phone. The call is established between user B and user C.
6. User B presses the hookflash again. MSAN_2 sends the hookflash signal to MSAN_1 through RFC2833. MSAN_1 informs the PBX of user B's hookflash pressing.
7. After receiving the hookflash message, the PBX transparently transmits the dial tone to user B. After hearing the dial tone, user B starts to dial the DTMF number 3. The dialed number is transparently transmitted to the PBX. By now, the 3WC is established between users A, B, and C.

NOTE
During the time from user B pressing the hookflash for the first time to user B pressing the hookflash for the second time and dialing the DTMF number 3, user A is in the call waiting state.

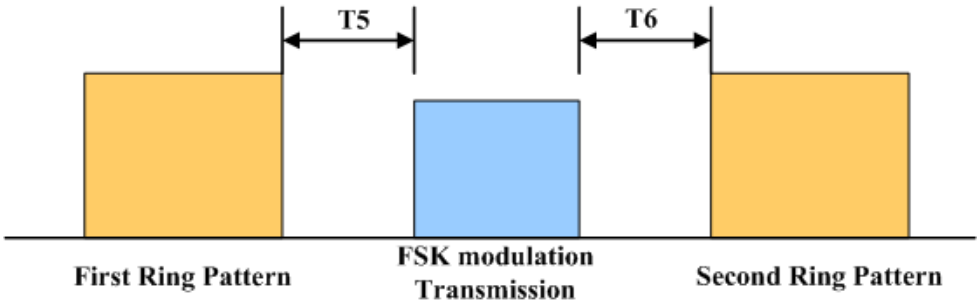
23.13.5 Ringing and CLIP Services for IP Z Interface Extension Feature

When a PSTN network is restructured into an IP network, the ringing and calling line identification presentation (CLIP) services for IP Z interface extension users need some implementation changes accordingly. This topic describes the context and implementation changes for the FXO port to support the ringing and CLIP services for IP Z interface extension users serving as the called parties.

Context

- Figure 23-85 illustrates the intervals of the ringing and CLIP signals in the on-hook process specified in ETSI EN 300 659-1.

  Figure 23-85 Intervals of the ringing and CLIP signals in the on-hook process (ETSI EN 300 659-1)

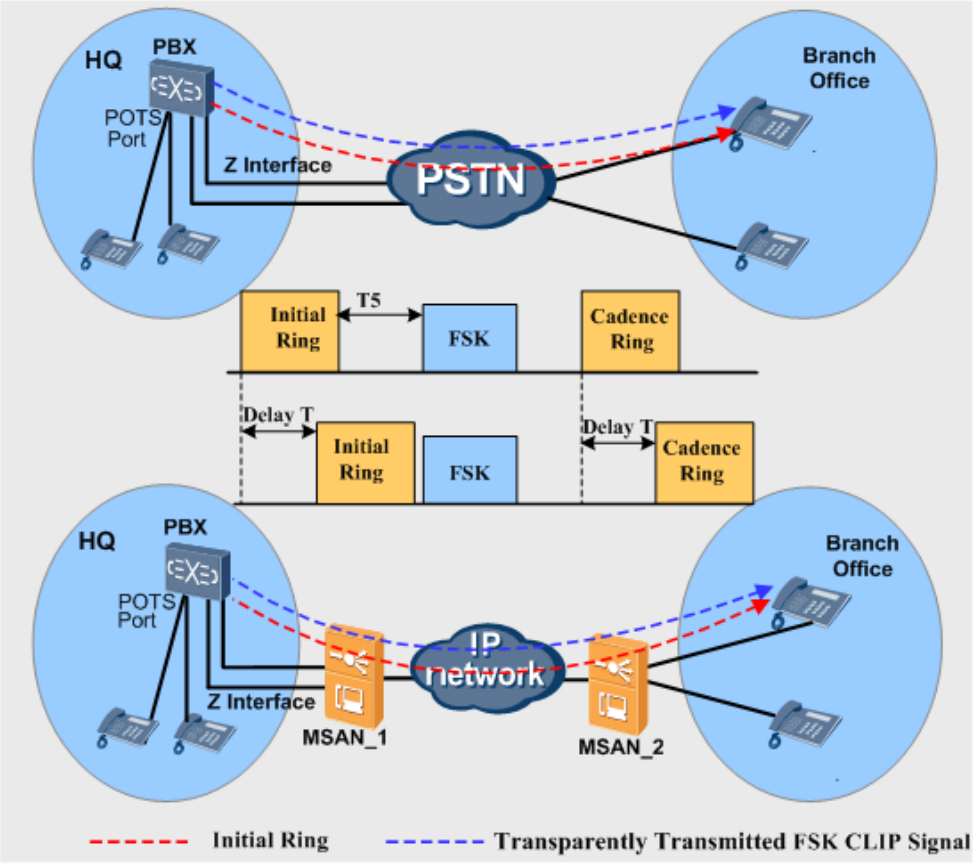
  According to the preceding figure, the interval between the first ring pattern signal and the FSK modulation transmission signal (CLIP signal) is T5 (500 ms ≤ T5 ≤ 2000 ms); the interval between the FSK modulation transmission signal and the second ring pattern signal is T6.

  The **First Ring** parameter in the protocol is the **Initial Ring** parameter of the MSAN, and the **Second Ring** parameter in the protocol is the **Cadence Ring** parameter of the MSAN.

- Compared with a PSTN network, an IP network needs more time (delay T) to send ringing signals of a Z interface extension call from the PBX to the telephone, because the ringing signals require processing by the MSAN on an IP network, as shown in Figure 23-86.

  However, both PSTN and IP networks transparently transmit FSK modulation transmission signals from the PBX to the telephone and thereby spend the same time. As a result, an IP network has a shorter T5 (interval between the first ring pattern signal and the FSK modulation transmission signal) than a PSTN network does.

Figure 23-86 T5 difference between a PSTN network and an IP network



Theoretically, when a call to an IP Z interface extension user is ended before it is answered, the PBX hook on then stop sending a ringing message to the MSAN_1 and then the MSAN_1 detect ringing missing message to send a ringing stop message to the MSAN_2. However, the ringing stop message is also delayed due to MSAN processing.

If the delay is not properly handled, the CLIP service may be affected, and the telephone may still ring after a call is ended.

Implementation Changes

- According to the preceding analysis, delay T must be shorter than T5 to ensure that the CLIP service is normal.

  Delay T = Ringing signal detection time of the FXO board + Time used by Z interface extension (about 150 ms, including codec, jitter buffer, and other message processing time)

  The ringing signal detection time of the FXO board can be configured using the **min-ontime** parameter of the **fxoport attribute set** command.

  The value of the **min-ontime** parameter must meet the following requirement: **min-ontime** < T5-150 ms.

  NOTE
  For the value of T5, see ETSI EN 300 659-1.

- The FXO board must spend as little time as possible in detecting ringing missing message to avoid the undesired telephone ringing after a call is ended.

The ringing missing signal detection time of the FXO board can be configured using the **max-offtime** parameter of the **fxoport attribute set** command.

NOTE

When you configure the **max-offtime** parameter, take the ringing break-make ratio into consideration. If the **max-offtime** parameter is set to a time shorter than the ringing break, ringing signals cannot be detected.

23.14 Key Techniques for Improving Voice Service Quality

Voice service quality is the biggest challenge faced by the IP telephony technology. IP telephony service has a higher requirements on real-time transmission of IP packets. If IP packets are lost, or transmission delay or jitter is introduced due to transmission errors or network congestion, subscribers hear noises during calls, and even more, ongoing calls may be interrupted. The VoIP technology provides a series of technologies, such as codec, echo cancellation (EC), and voice activity detector (VAD) to improve the voice quality.

23.14.1 Codec and Packetization Duration

Introduction

Codec is a key technology of voice services. Coding means that the DSP encodes the TDM-based voice data, assembles the data into packets, and then sends the packets to the IP network. Decoding means that the DSP decodes the voice packets received from the IP network and plays the voice to the TDM side.

Frequently-used codec types are G.711A, G.711Mu, G.729a, G.729b, G.726, G.723.1Low, and G.723.1High. G.711A and G.711Mu are lossless coding schemes. G.729, G.723.1Low, and G.723.1High are lossy compressed coding schemes. The compressed coding schemes require less bandwidth, but the voice quality is poor and the delay is large. (G.711 delivers the best voice quality but requires a bandwidth of 64 kbit/s. G.723 requires less bandwidth but the voice quality is less satisfying.)

PTime is the interval at which the DSP assembles the voice data into packets. It varies according to the codec type. Table 23-16 lists the codec types.

Table 23-16 Codec list

Codec Type	Coding Rate (kbit/s)	PTime
G.711A/G.711Mu	64	10 ms, 20 ms, 30 ms, 40 ms, 50 ms, or 60 ms
G.729a/G.729b	8	10 ms, 20 ms, 30 ms, 40 ms, 50 ms, or 60 ms
G.726	16/24/32/48	10 ms or 20 ms
G.723.1High	6.3	30 ms or 60 ms
G.723.1Low	5.3	30 ms or 60 ms

Standards and Protocols Compliance

ITU-T G.711, ITU-T G.729, and ITU-T G.723

23.14.2 EC

Introduction

Echo canceller (EC) is to cancel echo during calls. Echo is classified into the acoustic echo and electrical echo.

- Acoustic echo

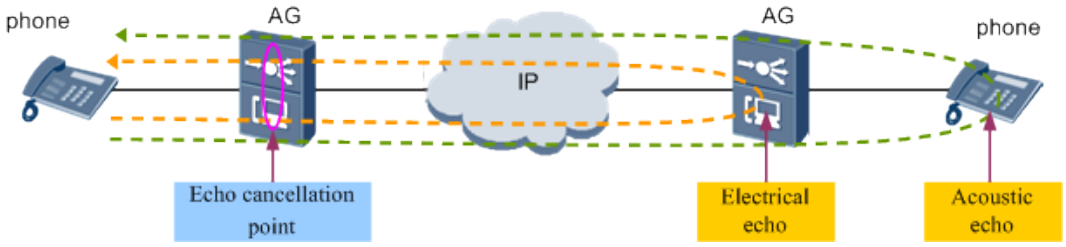
  Acoustic echo refers to the echo reflected by an obstacle when the voice encounters the obstacle in the transmission path. For example, if you place the phone at one side and speak at the other side, you can hear your own voice. This is because the voice is transmitted through the table and reflected from the collector to the receiver of the phone. Currently, the VoIP DSP chip does not support cancellation of the acoustic echo because it cannot distinguish the normal voice from the acoustic echo.

- Electrical echo

  Electrical echo is generated by the 2-wire/4-wire converter on the service board, because the impedance matching is not ideal on the 2-wire/4-wire converter. EC generally refers to the cancellation of the electrical echo.

Figure 23-87 shows how the electrical echo is generated.

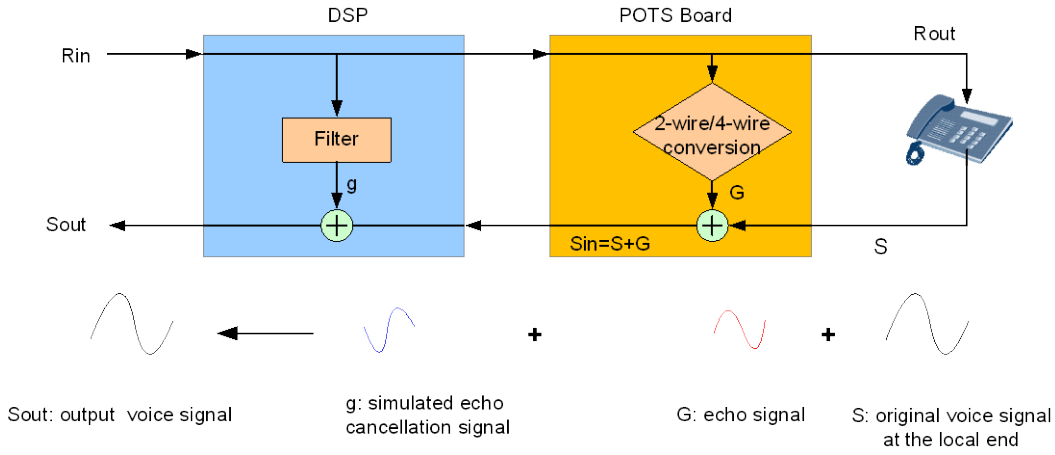
Figure 23-87 EC



In the PSTN network, owing to the small delay, the voice and the echo reach the ears of the speaker almost at the same time. Therefore, the echo can hardly be perceived. On the VoIP network, owing to the large delay, the echo reaches the ears some time after the voice is heard. Therefore, the echo can be easily perceived. As described in ITU-T G.131 and ITU-T G.161, the echo can be perceived when the echo delay exceeds 25 ms.

Figure 23-88 shows how the EC is implemented.

Figure 23-88 Implementation of the EC function



Rin is the voice received from the remote end. **Rin** is the input of the wave filter and the output of the wave filter is the simulated echo cancellation signal **g**. During the 2-wire/4-wire conversion, echo **G** is generated based on **Rin**. **S** is the original voice signal at the local end, that is, the voice received by the local receiver. The local-end voice signal **S** is overlaid with the echo cancellation signal **G**, resulting in the input signal of the EC, **Sin**. The EC removes the simulated echo **g** from the input signal **Sin** to obtain the output signal **Sout**.

Sin = S + G

Sout = Sin - g = S + G - g

G ≈ g

Therefore, Sout ≈ S

Reference Standards and Protocols

ITU-T G.168, ITU-T G.131, and ITU-T G.161

23.14.3 Non-Linear Processor

Introduction

Owing to various reasons, the EC cannot cancel all the echoes. To improve the EC performance, a non-linear processing (NLP) is performed on the remaining echoes when the power of the remaining echoes is lower than a preset value. This can further reduce the power of the remaining echoes. A simple method is to replace the remaining echoes with the silence when the power of the remaining echoes is lower than the threshold.

Specifications

The NLP function can be enabled or disabled by configuring the DSP profile on a port. If the DSP profile is not configured, the system automatically enables or disables the NLP function according to the service mode. Specifically, for the voice service, the system enables the NLP function; for the fax or modem service, the system disables the NLP function.

Impact

The NLP function must be disabled in the case of FoIP or MoIP.

Reference Standards and Protocols

ITU-T G0.168, ITU-T G0.131, and ITU-T G0.161

23.14.4 VAD/CNG

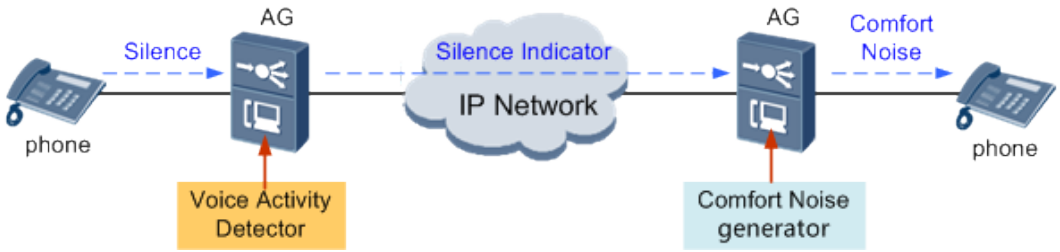
Introduction

According to statistics, silent duration exceeds 50% of the total session duration. If data is still transmitting in common packetization mode during the silent period, network bandwidth resources will be wasted. The voice activity detector (VAD) and comfort noise generator (CNG) significantly reduce network bandwidth usage to ease the insufficiency of network resources.

- VAD: The VAD identifies voice and silent durations based on signal energy. After detecting silence, the Tx end only sends mute indication packets instead of voice IP packets, lowering occupied bandwidth.
- CNG: To avoid long time silence during the mute period that may discomfort the user, the Rx end needs to generate comfort noises during the mute period according to the mute indication sent by the Tx end.

VAD and CNG are always used together. VAD is used on the Tx end, and CNG is used on the Rx end. For example, when VAD is enabled, the DSP packetizes RTP packets and sends the packets to the remote end only when it detects voice signals. In the case of silence, the DSP does not send RTP packets to the IP side. The DSP sends a silence ID (SID) to inform the remote end only when the background noise changes. The remote DSP then generates background noises according to the information carried in the SID. Figure 23-89 shows the implementation process.

Figure 23-89 Networking for VAD and CNG



Specifications

ITU-T G.711 and ITU-T G.729

23.14.5 PLC

Introduction

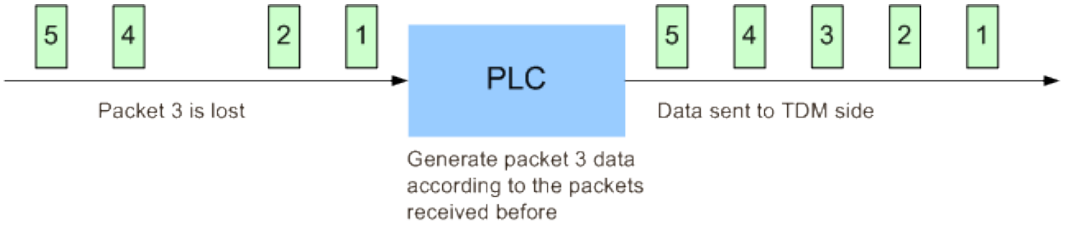
When a network or a device loses packets, the voice quality deteriorates. In practice, packet loss is inevitable. If the packet loss concealment (PLC) is enabled to compensate the signals, however, the impact of packet loss on the voice quality is reduced and the success rates of FoIP and MoIP services increases in the case of packet loss. Voice IP data is transmitted using Real-Time Transport Protocol (RTP). The DSP chip detects the sequence number of received RTP packets. If the DSP chip detects packet loss, it automatically constructs voice data based on a configured compensation algorithm and sends the data to the TDM side.

Three compensation modes are available:

- Compensate the lost packet with the silence.
- Compensate the lost packet with the previous packet.
- Compensate the lost packet with a similar packet that is calculated based on the energies of the previous packet and the subsequent packet (as described in G.711 Appendix I).

The third mode consumes the most DSP resources, but improves the voice quality in the most satisfying manner. The first mode consumes the least DSP resources, but improves the voice quality in the least satisfying manner. Figure 23-90 shows the packet loss compensation using the previous packet.

Figure 23-90 Packet loss compensation using the previous packet



NOTE
When the IP network quality is poor, if the number of consecutively lost RTP packets is greater than 2, the voice quality is still poor even if the PLC is enabled.

Reference Standards and Protocols

G.711 Appendix I

23.14.6 JB

Introduction

The transmission quality on the IP network is not guaranteed. The interval at which packets are received from the remote end is not even, and the sequence of packets received may be different from the sequence that these packets are sent. As a result, the voice quality is degraded. Therefore, the jitter buffer (JB) is introduced to eliminate the jitter of the IP network. The basic idea of JB is that delay is introduced so that uneven and disordered RTP packets can be sequenced in the buffer and then sent to the TDM side. Figure 23-91 shows the JB implementation method.

The JB is classified into the dynamic JB and the static JB.

- Dynamic JB: The buffer depth can be automatically adjusted based on network jitter conditions, so that the introduced delay is appropriate to process the jitter. The dynamic JB mainly applies to voice services.
- Static JB: The buffer depth is fixed and cannot be changed based on network jitter conditions. The static JB mainly applies to data services, such as fax over IP (FoIP) and modem over IP (MoIP). This is because the modification of buffer depth may lead to packet loss, which has great adverse impact on data services.

Figure 23-91 JB implementation method

Specifications

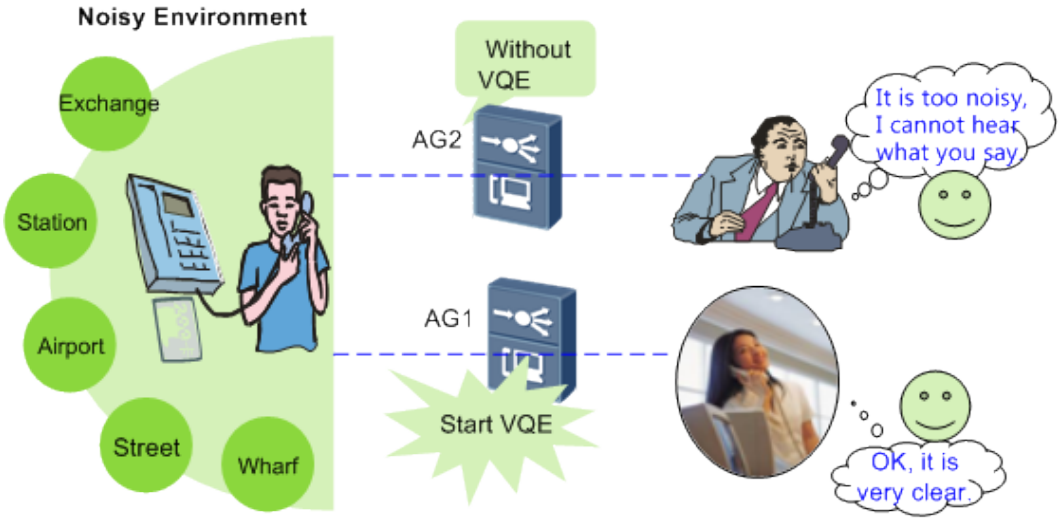
The dynamic JB and the static JB are supported. The adjustable range of the JB depth is 0 ms to 135 ms.

23.14.7 VQE

Introduction

The Voice Quality Enhanced (VQE) feature applies to voice services in the noisy public areas, such as the roads, docks, scenic spots, and bus stations. Deployment of VQE in these areas can improve the voice quality and user experience. Figure 23-92 shows application scenarios and advantage of the VQE.

Figure 23-92 Application scenarios and advantage of the VQE



The VQE consists of two functions, automatic gain control (AGC) and spectral noise suppression (SNS).

- AGC: A target gain energy value is configured so that the output gain can be automatically adjusted during the VoIP communication, ensuring that subscriber can still hear voices in a noisy environment. AGC ensures smooth energy adjustment and prevents adverse impact brought by abrupt energy changes.
- SNS: A target noise suppression value is configured. After detecting that the noise energy value is greater than the target value, AG reduces the noise energy value to enable subscribers to feel comfortable during VoIP communication.

Specifications

- The AG only supports AGC. The VQE feature is configured based on user ports and takes effect for calls initiated after the VQE is enabled.
- The AG supports VQE only using the G.711 codec. If a user configures the VQE feature using non-G.711 codecs, the VQE does not take effect, and the AG does not display any information.

23.14.8 Fax/Modem Quality Enhancement

Overview

After the IP network takes the place of the PSTN network, the use of fax and modem on the VoIP network becomes more and more popular. Therefore, the AG is required to provide applications similar to those of the PSTN network. Currently, the Voice Band Data (VBD) transparent transmission is adopted by the medium gateway (MG) in the application of the fax and modem. Transparent transmission, however, relies heavily on the bearer network and deterioration of the network quality may lead to service failures.

The fax/modem quality enhancement feature consists of the RFC2198 intelligent startup function and the packetization at the interval of 10 ms. After the fax/modem quality enhancement feature is enabled, the RFC2198 function and the packetization at the interval of 10 ms are automatically started. The following table provides more details.

Problem	Solution	Description
Packet loss	RFC 2198	The RFC 2198 standard uses the data stream redundancy mechanism to prevent the packet loss of the network from degrading the service quality. When the average consecutive packet loss ratio is low, the receiver can reassemble and restore the lost packet based on the redundant packets in the later received packets. The audio redundancy mechanism described in RF2198 can be used to restore the events lost in the packets, while the mode described in RFC2833 can be used to process the DTMF signals transmitted through the RTP packets.
Network delay	10-ms packetization	Information carried by packets assembled every 10 ms is less the information carried by packets assembled every 20 ms. Therefore, in case of packet loss, the packetization at the interval of 10 ms causes less impact on services that the packetization at the interval of 20 ms. Using this technology, the device automatically detects fax and modem signals and switches the 20-ms packetization interval (intended for voice services) to the 10-ms packetization

Problem	Solution	Description
		interval, thereby reducing the network delay of fax and modem data transmission.

The enhanced quality feature of the fax and modem is mainly used to improve the put-through rate and online duration of the fax and modem services. For example, if POS terminals are connected to a modem in a shopping mall or a bank, the fax/modem quality enhancement feature can be used to improve the stability and online duration of the Modem, thus preventing the disconnection caused by the poor network quality, as shown in Figure 23-93.

Figure 23-93 Enhanced Modem

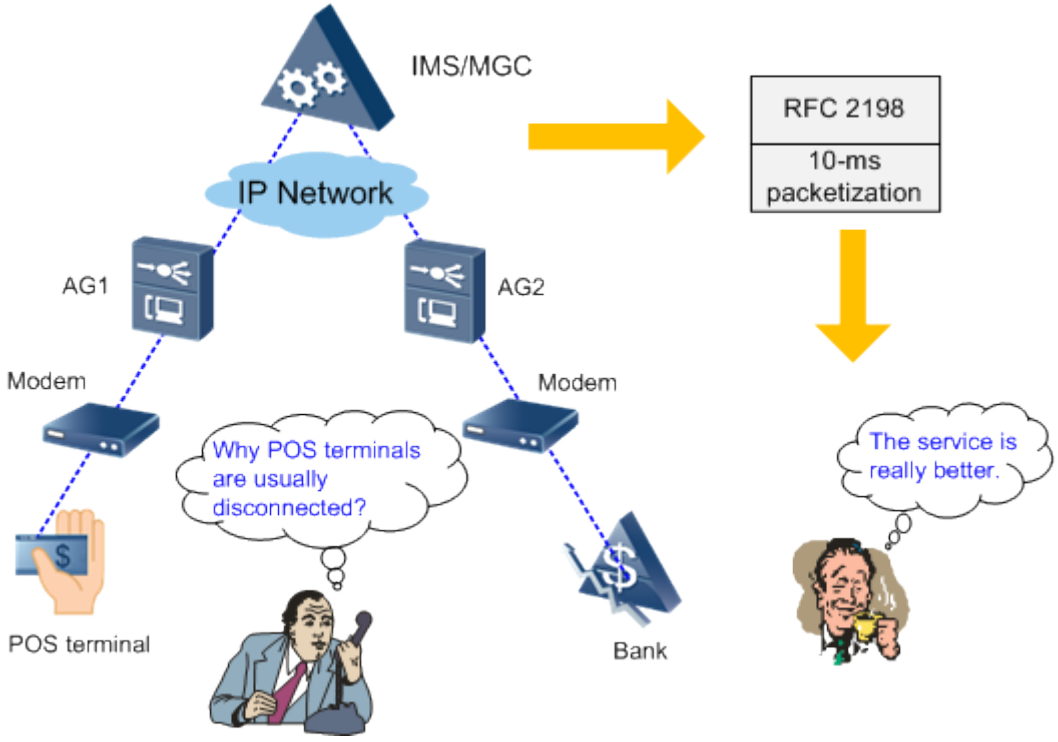


Table 23-17 lists the test data recorded before and after the quality enhancement feature is enabled. The data shows that stability of services improves substantially after the quality enhancement feature is enabled.

Table 23-17 Test data before and after the quality enhancement

Modem	PTime	Network Packet Loss Ratio (Random)	Online Duration
T336CX	20 ms	0.10%	22 hours
		0.50%	1.5 hours
		1.00%	Unavailable

Modem	PTime	Network Packet Loss Ratio (Random)	Online Duration
		1.00% (rfc2198)	12.5 hours
	10 ms	0.10%	> 24 hours
		0.50%	22.5 hours
		1.00%	5.5 hours
		1.00% (rfc2198)	24 hours

Specifications

The fax/modem quality enhancement is supported.

23.15 Voice Service Maintenance and Diagnosis

The maintenance and diagnosis features of voice services include these features such as the loop-line test, circuit test, call emulation test, continuity test, VBD fault diagnosis, Real-time Transport Control Protocol (RTCP) statistics and so on.

23.15.1 Call Emulation Test

A call emulation test emulates call functions to verify data configuration for the voice service. The call emulation test can also be used to locate voice service faults.

Introduction to the Call Emulation Test

Definition

In a call emulation test, the device emulates the call function of a voice user. It is used to test the services on the POTS user port. A call emulation test includes:

- **Calling party emulation test**: The POTS user port on the device functions as the calling party. In this test, a test engineer acting as a called party is required.
- **Called party emulation test**: The POTS user port on the device functions as the called party. In this test, a test engineer acting as a calling party is required.
- **Calling and called party emulation test**: Two POTS user ports function as the calling and called parties. The test does not require manual operations. Specifically, the device automatically performs the process from calling party off-hook to called party off-hook to set up a call between the two parties and stops the test after the call hold time expires.

Application Scenarios

A call emulation test can be used in the following scenarios:

- Acceptance test during a new deployment: The software and hardware functions, including service configurations, of the device need to be verified after the device is installed. The verification ensures follow-up service provisioning.

Traditionally, the engineer goes to the device installation site, makes cables, connects a test phone set to the ONT, uses the test phone set as a caller or callee, and verifies basic voice services.

- Fault locating in the OAM phase: After the device enters the OAM phase, it is usually necessary to test basic voice services in order to locate a fault. In the access network, however, a large number of devices are installed in complicated environment, geographically dispersed, and remotely located. It is inconvenient and costly either to test a newly installed device or locate faults.

Call emulation tests can be conducted remotely. In a call emulation test, the test engineer does not need to prepare cables on the device installation site, connect test terminals to the device, or perform dialup tests on site. Instead, the test engineer enables the call emulation function in the maintenance center through the command line interface (CLI) or network management system (NMS) and uses a test phone set in the central office (CO) to make calls to the emulation port on the device. In this way, the test engineer can verify the data configurations and basic service functions.

Benefits

- The call emulation feature can be used to remotely verify and accept services and locate faults, which greatly reduces the operating expense (OPEX) for carriers.
- The call emulation feature shortens fault location time, which significantly improves the fault locating efficiency.

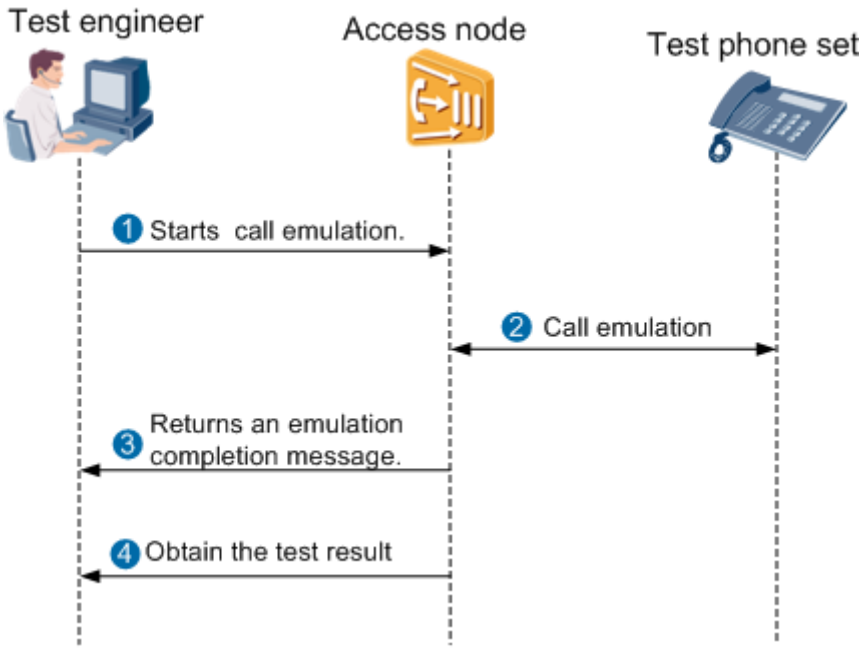
Principles of the Call Emulation Test

In a call emulation test, the system emulates the actions of a user port to achieve the emulation of the calling party and called party. The actions of a user port include off-hook, on-hook, number dialing, and ringing detection. Figure 23-94 shows the principles of a call emulation test. The test phone set, placed at the central office (CO), is used to perform a call emulation test with a configured port on the access node. This test method remotely checks whether the voice service on the device is functional.

NOTE

For a calling and called party emulation test, no phone set is required. Both the calling and called parties are emulated by the POTS ports on the access node.

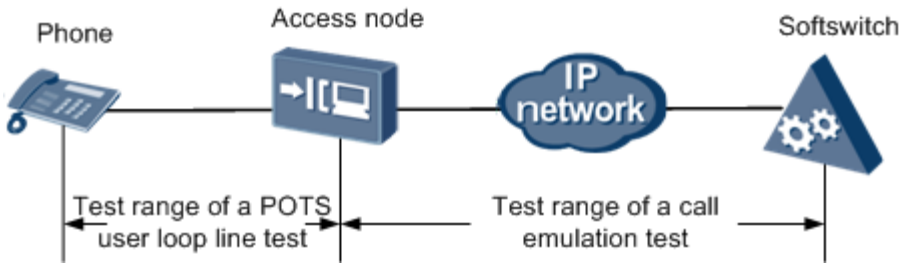
Figure 23-94 Principles of a call emulation test



Networking Application

Figure 23-95 shows the example network of a call emulation test.

Figure 23-95 Example network of a call emulation test



A call emulation test checks whether the voice service is functional only on the network side of the device.

- If a conversation channel cannot be established during a call emulation test, check whether the network data configurations are correct.
- If the conversation channel is established but the voice is not clear during a call emulation test, check whether the cables are connected properly.

If test engineers need to check whether the voice service is functional on the user side, they can:

- Perform a POTS user loop line test to check whether the line between the device and the user phone set is functional. For details, see 23.15.2 POTS User Loop Line Test.

- Perform a POTS user circuit test to check whether the POTS board on the device is functional. For details, see 23.15.3 POTS User Circuit Test.

Calling Party Emulation Test

In a calling party emulation test, the device port emulates user off-hook, number dialing, communication, and on-hook.

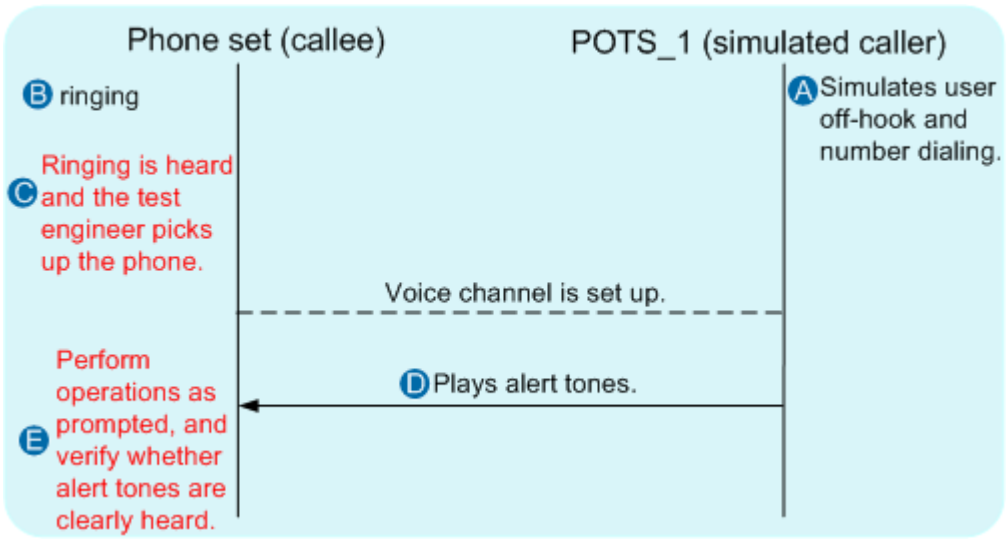
1. Start a calling party emulation test.

   On the access node, the test engineer sets the POTS_1 port as the calling party emulation port, configures the phone number to be dialed, and starts the calling party emulation test. If the function of playing alert tones upon the test beginning is enabled, alert tones are played after the test is started.

2. Initiate the calling party emulation test.

   The information marked red in the following figure indicates the operations that need to be performed by the test engineer on the test phone set at the CO.

   Figure 23-96 Interaction between the POTS_1 port and phone set during a calling party emulation test



   a. The calling party emulation test is started on the access node and the calling party emulation port automatically emulates user off-hook. After detecting the dial tone, the port emulates number dialing.

   b. The called party (whose number is automatically dialed by the calling party emulation port) waits for the phone to ring. If the phone rings, the signaling channel is functional and the configured data is correct. If the phone does not ring, the test engineer needs to check service data (such as route, VLAN, and core-network data), troubleshoot the voice fault if any, and perform the test again.

   c. The called party picks up the phone. The call is set up.

   d. The calling party emulation port plays announcements for the called party.

   e. The test engineer checks whether the announcements are clearly heard. After hearing the announcements, the test engineer presses the specified verification number (a matched DTMF number, the asterisk key (*) by default), indicating that the media channel is functional. The test result is "Test Succeed."

NOTE

If the function of playing alert tones based on the DTMF matching result is enabled, the system plays the alert tones after the test engineer presses the DTMF number. The alert tones include the DTMF number matching success alert tone and DTMF number matching failure alert tone.

3. Obtain the test result. If the test fails, the system outputs the failure causes, with which the test engineer can identify the possible cause of the fault. Table 23-18 lists the test results.

Called Party Emulation Test

In a called party emulation test, the POTS port on the access node emulates a called party. The called party emulation port automatically emulates user off-hook after detecting the ringing current.

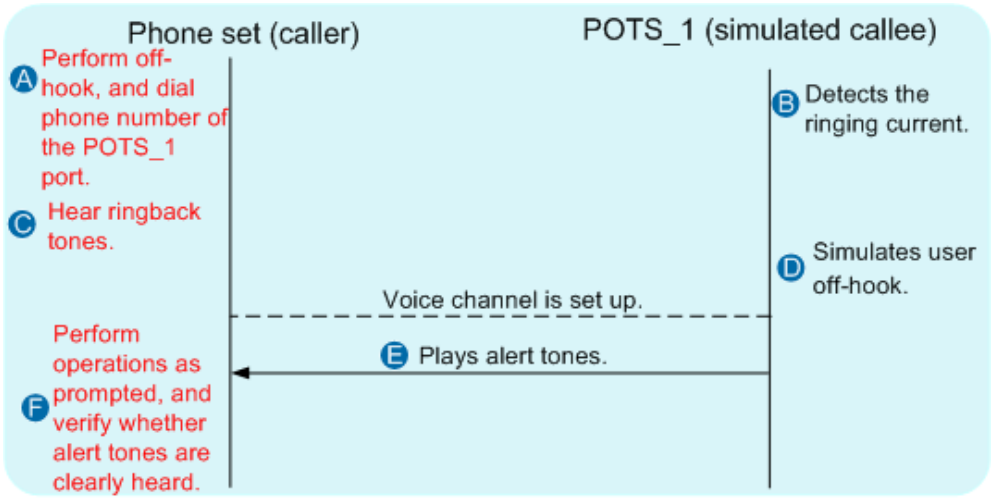
1. Start a called party emulation test.

   On the access node, the test engineer sets the POTS_1 port as the called party emulation port and starts the called party emulation test. If the function of playing alert tones upon the test beginning is enabled, alert tones are played after the test is started.

2. Initiate the called party emulation test.

   The information marked red in the following figure indicates the operations that need to be performed by the test engineer on the test phone set at the CO.

   Figure 23-97 Interaction between the POTS_1 port and phone set during a called party emulation test



   a. The test engineer picks up the phone, hears the dial tone, and dials the number of the POTS_1 port.

   b. If the called party emulation port of the device detects the ringing current, the configured user data is correct. If the called party emulation port does not detect the ringing current, the test engineer needs to check service data (such as route, VLAN, and core-network data), troubleshoot the voice fault if any, and perform the test again.

   c. If the calling party hears ringback tones before the called party picks up the phone (in this test, off-hook is automatically emulated by the called party emulation port), the signaling channel is functional. Otherwise, the test engineer needs to verify the configured service data and perform the test again.

d. The called party emulation port emulates off-hook. The call is set up.

e. The called party emulation port plays announcements for the calling party.

f. The test engineer checks whether the announcements are clearly heard. After hearing the announcements, the test engineer presses the specified verification number (a matched DTMF number, the asterisk key (*) by default), indicating that the media channel is functional. The test result is "Test Succeed."

NOTE

If the function of playing alert tones based on the DTMF matching result is enabled, the system plays the alert tones after the test engineer presses the DTMF number. The alert tones include the DTMF number matching success alert tone and DTMF number matching failure alert tone.

3. Obtain the test result. If the test fails, the system outputs the failure causes, with which the test engineer can identify the possible cause of the fault. Table 23-19 lists the test results.

Calling and Called Party Emulation Test

In a calling and called party emulation test, two POTS ports emulate the calling party and called party respectively. The calling party emulation port emulates the actions of the calling party, while the called party emulation port emulates the actions of the called party. These two ports automatically emulate calling party off-hook, number dialing, called party off-hook (after detecting the ringing current), mutual DTMF number sending for media channel verification, and on-hook after the verification.

NOTE

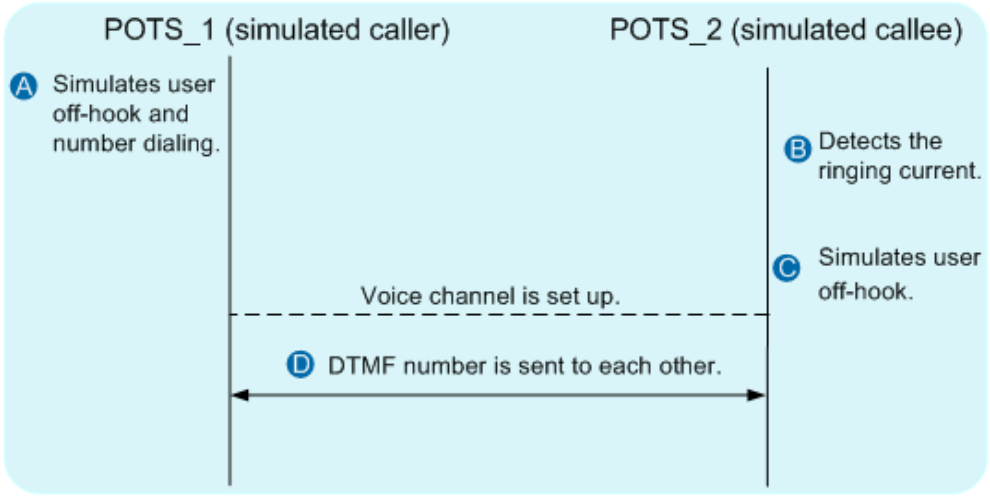
A calling and called party emulation test does not require any manual operation, which greatly improves the test efficiency. In this test, the device automatically verifies whether the media channel is functional. However, the device's sensitivity to media streams may vary from the human ears' sensitivity to media streams. Therefore, the call quality cannot be verified.

1. Start a calling and called party emulation test.

   On the access node, the test engineer sets the POTS_1 port as the calling party emulation port and the POTS_2 port as the called party emulation port, configures the phone number to be dialed, and starts the calling and called party emulation test.

2. Initiate a calling and called party emulation test.

Figure 23-98 Interaction between the POTS_1 and POTS_2 ports during a calling and called party emulation test



a. The calling and called party emulation test is started on the access node and the POTS_1 port automatically emulates user off-hook. After detecting the dial tone, the port emulates number dialing.

b. The POTS_2 port (whose number is automatically dialed by the POTS_1 port) waits and checks whether the ringing current can be detected. If the POTS_2 port detects the ringing current, the signaling channel is functional and the configured user data is correct. If the POTS_2 port does not detect the ringing current, the test engineer needs to check service data (such as route, VLAN, and core-network data), troubleshoot the voice fault if any, and perform the test again.

c. The POTS_2 port emulates off-hook. The call is set up.

d. The POTS_1 and POTS_2 ports send the DTMF number to each other to verify whether the media channel is functional. If the DTMF numbers sent by POTS_1 and POTS_2 ports are correct, the media channel is functional. The test result is "Test Succeed."

3. Obtain the test result. If the test fails, the system outputs the failure causes, with which the test engineer can identify the possible cause of the fault. Table 23-18 and Table 23-19 list the test results.

Test Results

Table 23-18 Results of a calling party emulation test

Test Result	Description
Test Succeed	The test is successful. The media channel of the test port is functional.
Test Failed: No dialing tone is played when the calling party dials a number	The calling party emulation port does not detect the dial tone. Possible causes are as follows: • The calling party emulation port does not detect the off-hook signal. • The off-hook signal is not reported. • No dial tone is issued after the off-hook signal is reported.
Test Failed: Busy tone is played after the calling party picks up the phone off the hook	The calling party emulation port detects the busy tone. Possible causes: The calling party emulation port does not subscribe to the POTS services or the digital signal processor (DSP) is faulty.
Test Failed: Busy tone is played when the calling party dials a number	The calling party emulation port detects the busy tone during number dialing. Possible cause: The number that the calling party emulation port dials does not match the digitmap.
Test Failed: The calling party does not dial a number	The calling party emulation port does not automatically emulate number dialing after detecting the dial tone. Possible cause: The internal processing mechanism of the system encounters an error.

Test Result	Description
Test Failed: Busy tone is played after the calling party dials a number	The calling party emulation port detects the busy tone after number dialing. Possible causes: The dialed number is busy or the dialed number is incorrect.
Test Failed: The calling party does not communicate with the called party after dialing a number	The call is not set up. Possible cause: The called party does not pick up the phone.
Test Failed: Release before pick-up of the calling party	The call is released before the calling party and called party enter a conversation. Possible cause: The signaling processing mechanism encounters an error.
Test Failed: The number matching of the calling party is not complete	The calling party emulation port fails to match the DTMF number sent by the called party after entering the conversation. Possible causes: The called party does not press the DTMF number or the DTMF number is lost during the transmission.
Test Failed: The number sending of the calling party is not complete	The called party does not complete the sending of the DTMF number to the calling party emulation port after entering the conversation. Possible cause: The called party hangs up the phone before the sending of the DTMF number is completed.
Test Failed: The number matching of the calling party fails	The DTMF number sent from the called party is incorrect.
Test Failed: The calling port is abnormal	The calling party emulation port is faulty. Possible causes are as follows: • The board is faulty. • The board is removed. • The calling party emulation port is faulty.

Table 23-19 Results of a called party emulation test

Test Result	Description
Test Failed: The phone of the called party does not ring	The called party emulation port does not receive any call. Possible causes: The calling party dials the wrong number or the signaling transmission encounters an error.
Test Failed: The called party does not pick up the phone off the hook	The called party emulation port does not automatically emulate off-hook after detecting the ringing current. Possible cause: The internal processing mechanism of the system encounters an error.

Test Result	Description
Test Failed: Busy tone is played after the called party picks up the phone off the hook	The called party emulation port detects the busy tone after off-hook. Possible cause: The calling party hangs up the phone.
Test Failed: The called party does not communicate with the calling party	The called party emulation port cannot enter the conversation after off-hook. Possible cause: The called party emulation port emulates off-hook so slowly that the calling party has hung up the phone.
Test Failed: Release before pick-up of the called party	The call is released before the calling party and called party enter a conversation. Possible cause: The signaling processing mechanism encounters an error.
Test Failed: The number matching of the called party is not complete	The called party emulation port fails to match the DTMF number sent by the calling party after entering the conversation. Possible causes: The calling party does not press the DTMF number or the DTMF number is lost during the transmission.
Test Failed: The number sending of the called party is not complete	The calling party does not complete the sending of the DTMF number to the called party emulation port after entering the conversation. Possible cause: The calling party hangs up the phone before the sending of DTMF number is completed.
Test Failed: The number matching of the called party fails	The DTMF number sent from the calling party is incorrect.
Test Failed: The called port is abnormal	The called party emulation port is faulty. Possible causes are as follows: • The board is faulty. • The board is removed. • The called party emulation port is faulty.

Configuring the Call Emulation Test

Prerequisites

- The voice service must be configured.
- A normal phone must be provided.

Procedure

Run the **simulate call parameter** command to configure the parameters for the call emulation test.

The type of the call emulation alert tone and DTMF number must be set before the automatic call emulation test is performed on the port.

By default, the type of the call emulation alert tone is voice announcement.

Step 1 Run the **display simulation call parameter** command to query the parameters configured for the current call emulation test.

Step 2 Start a call emulation test. You can start a calling party emulation test, a called party emulation test, or a calling and called party emulation test based on actual requirements.

- To start a calling party emulation test, run the **simulate call start caller** command.
- To start a called party emulation test, run the **simulate call start callee** command.
- To start a calling and called party emulation test, run the **simulate call start call** command.

----End

Result

After completing the call emulation test, the device directly outputs the test result. The test result is used to check whether the call is normal.

23.15.2 POTS User Loop Line Test

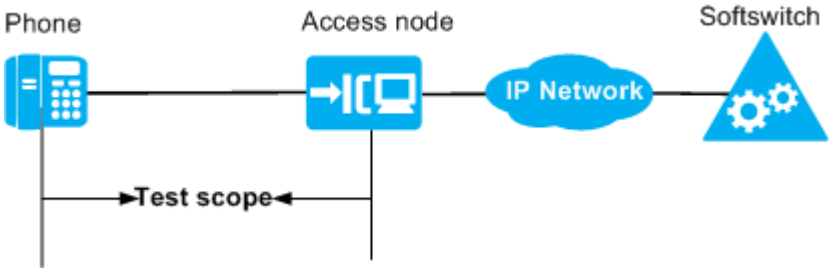
A POTS user loop line test is used to test the electrical indicators of the line from the test device (an access node) to a phone. When users' POTS services are faulty, POTS user loop line tests can be performed to test the performance and electrical indicators of the loop line to diagnose whether the loop line is faulty.

NOTE
The to-be-tested POTS port must not be faulty.

Networking Application

Figure 23-99 shows the example network of a POTS user loop line test. A POTS user loop line test is used to locate faults on the line from an access node to a phone.

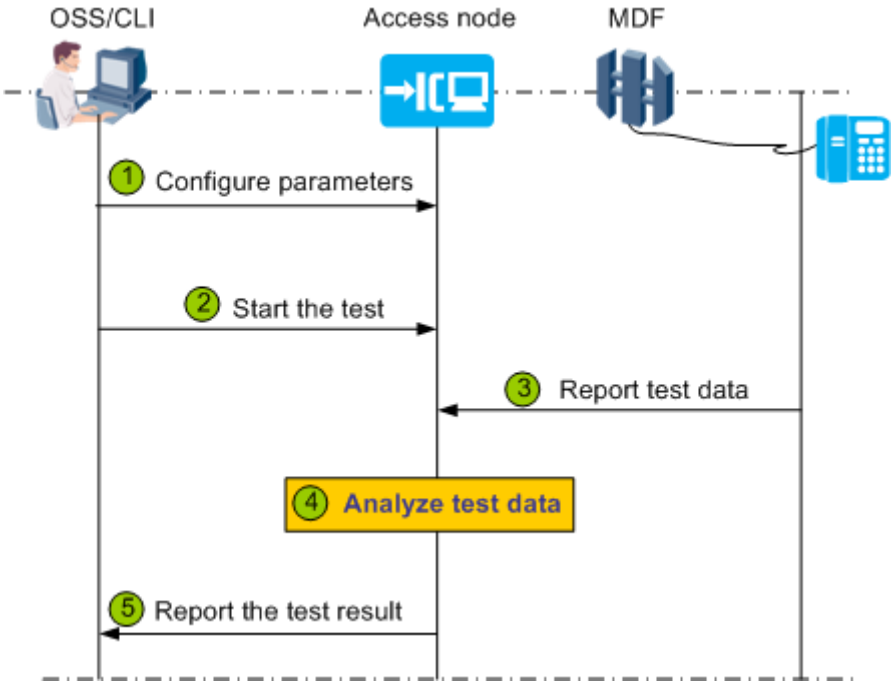
Figure 23-99 Example network of a POTS user loop line test



Test Procedure

Figure 23-100 shows the test procedure of a POTS user loop line test.

Figure 23-100 Test procedure of a POTS user loop line test



1. (Optional) Maintenance engineers set the parameters for the access node on the NMS or remotely log in to the access node from a management PC to set the parameters.

NOTE
Generally, the default parameter values are used, and maintenance engineers do not need to set them.

- In test mode, run the **pots test-para** command to set the physical layer parameters.

  During a loop line test, to avoid affecting the services and functions of the live network, it is necessary to set the physical layer parameters to control the electrical indicators, such as the maximum and minimum voltages supported by the test.
- In test mode, run the **pots loop-line-threshold** command to set thresholds of the test.

  The access node uses these thresholds as the criteria to check whether the line is faulty when analyzing the test data.

2. Maintenance engineers start a loop line test by using the NMS or running the **pots loop-line-test** command.

   If users (connected to the access node) are making calls during a loop line test, maintenance engineers can cancel, forcibly perform, or delay the test based on actual conditions.

NOTE
If maintenance engineers forcibly perform the test, services on the port are interrupted and users' telephone services are affected. Therefore, exercise caution when performing this operation.

3. The access node collects the data of the test items. Table 23-20 lists the loop line test items.

Table 23-20 Loop line test items

Test Item	Specific Test Item

Test Item	Specific Test Item
Voltage	A->G DC voltage
	B->G DC voltage
	A->B DC voltage
	A->G AC voltage
	B->G AC voltage
	A->B AC voltage
	A->G AC frequency
	B->G AC frequency
	A->B AC frequency
Resistance	A->ground insulation resistance
	B->ground insulation resistance
	A->B insulation resistance (low)
	B->A insulation resistance (low)
	A->B insulation resistance (high)
	B->A insulation resistance (high)
Capacitance	A->ground capacitance
	B->ground capacitance
	A->B capacitance (low)
	A->B capacitance (high)
Conductance	A->ground conductance
	B->ground conductance
	A->B conductance (low)
	A->B conductance (high)
Susceptance	A->ground susceptance
	B->ground susceptance
	A->B susceptance (low)
	A->B susceptance (high)
Current	A->ground DC current
	B->ground DC current
	A->B DC current
	B->A DC current

Test Item	Specific Test Item
	A->ground AC current
	B->ground AC current
	A->B AC current
	B->A AC current

4. The access node analyzes the collected data according to the algorithm, and outputs the test conclusion.
5. The access node reports the test conclusion, and maintenance engineers diagnose whether the tested line is faulty based on the test conclusion.

Test Conclusion

Table 23-21lists the loop line test conclusions.

Table 23-21 OLT loop line test conclusions

Item	Conclusion
Line state	Normal
	A->ground AC voltage is hazardous to persons
	B->ground AC voltage is hazardous to persons
	AB->ground AC voltage is hazardous to persons
	A->ground EMF AC voltage exist
	B->ground EMF AC voltage exist
	AB->ground EMF AC voltage exist
	A->ground abnormal AC voltage exist
	B->ground abnormal AC voltage exist
	AB->ground abnormal AC voltage exist
	A->ground DC voltage is hazardous to persons
	B->ground DC voltage is hazardous to persons
	AB->ground DC voltage is hazardous to persons
	A->ground EMF DC voltage exist
	B->ground EMF DC voltage exist
	AB->ground EMF DC voltage exist
	A->ground abnormal DC voltage exist
	B->ground abnormal DC voltage exist

Item	Conclusion
	AB->ground abnormal DC voltage exist
	A line grounding
	B line grounding
	AB line grounding
	A->ground resistance fault
	B->ground resistance fault
	AB->ground resistance fault
	A->ground resistance leak
	B->ground resistance leak
	AB->ground resistance leak
	AB->ground poor insulation
	AB->ground capacitance leak
	A->ground capacitance leak
	B->ground capacitance leak
	Double line break or no terminal
	Cut off in MDF (that is, a line cut occurs between the main distribution frame and the device)
	Cut off out MDF (that is, a line cut occurs between the main distribution frame and the user side)
	Self mixed in MDF (shorted wires within the same twisted pair, occurring between the main distribution frame and the device)
	Self mixed out MDF (shorted wires within the same twisted pair, occurring between the main distribution frame and the user side)
PPA test result **NOTE** A passive test termination (PPA) is similar to a test reference point. It is used to detect whether a fault occurs on the loop line between a point and the PPA so that maintenance engineers can locate faults section by section.	PPA not detected
	A->B PPA detected
	B->A PPA detected
	A->B 2 PPA detected
	B->A 2 PPA detected
Terminal status	Phone not connected
	Off hook

Item	Conclusion
	ETSI Signature or Elec ring circuit
	A-B short
	R-C network (on hook or modem exist)
	Electronic ringing circuit
	Other terminal

References

Reference standard and protocol: ITU-T G.996.2 Single-ended line testing for digital subscriber lines (DSL)

23.15.3 POTS User Circuit Test

A POTS user circuit test is used to check whether the chip of a POTS board functions normally. If the POTS services are faulty and the loop line works normally, POTS user circuit tests can be used to test the functions (such as the ringing and power feeding) and some parameters (such as the feeding voltage and ringing voltage) of the board circuit to check whether the circuit works normally.

Feature Dependency and Limitation

- The to-be-tested POTS port must not be faulty.
- Only one circuit test can be started on a POTS board at a time.

Test Procedure

1. Maintenance engineers start a circuit test by using the NMS or running the **pots circuit-test** command.

   If users are making calls during a circuit test, maintenance engineers can cancel, forcibly perform, or delay the test based on actual conditions. If maintenance engineers forcibly perform the test, services on the port are interrupted and users' telephone services are affected. Therefore, exercise caution when performing this operation.

   NOTE
   If maintenance engineers forcibly perform the test, services on the port are interrupted and users' telephone services are affected. Therefore, exercise caution when performing this operation.

2. Maintenance engineers check whether the circuit is faulty based on the test results. The results of a circuit test include: Normal, Abnormal, and Not supported. Table 23-22 lists the specific test items and exception description of a circuit test.

Table 23-22 Test items and exception description of a circuit test

Test Item	Exception Description
Digital voltage	Indicates the digital voltage. Users cannot make calls if this item is abnormal.
Low power supply voltage (negative)	Indicates the low-voltage power supply

Test Item	Exception Description
	(negative) to the POTS chip. The power consumption of the board increases if this item is abnormal.
High power supply voltage (negative)	Indicates the high-voltage power supply (negative) to the POTS chip. Users cannot make calls or the ringing tone is irregular if this item is abnormal.
Positive power supply voltage	Indicates the high-voltage power supply (positive) to the POTS chip. The ringing tone is irregular if this item is abnormal.
Off hook detective	Indicates the off-hook detection function of the POTS chip. Users cannot make calls if this item is abnormal.
On hook detective	Indicates the on-hook detection function of the POTS chip. Users cannot make calls if this item is abnormal.
A->B feeder voltage	Indicates the output voltage of the POTS chip. The call quality may be impaired if this item is abnormal.
A->ground feeder voltage	
B->ground feeder voltage	
A->B feeder voltage	Indicates the output voltage of the POTS chip. The call quality may be impaired if this item is abnormal.
Ringing current voltage	Indicates the output ringing voltage of the POTS chip. The ringing tone is excessively low if this item is abnormal.
Ringing current frequency	Indicates the output ringing frequency of the POTS chip. The ringing tone is irregular if this item is abnormal.
Stop ringing	Indicates the ringing stopping frequency of the POTS chip. Users can hear the ringing tone, but cannot communicate with the peer party after picking up the phone if this item is abnormal.
Loop current	Indicates the output current of the POTS chip. The call quality may be impaired if the output current is lower than 18 mA.

23.15.4 POTS Port Loop Test

A POTS port loop test is used to test the hardware and configurations related to POTS services during device installation or before POTS service provisioning. It helps reduce the number of site visits and minimize maintenance costs.

Overview

Maintenance engineers can locally start a POTS port loop test on the device, or remotely log in to the device and then start the test. A POTS port loop test consists of two parts:

- Device hardware test: This test targets at access nodes that are not yet provisioned with the voice service. When an access node is newly deployed, the hardware of the voice module needs to be tested to evaluate the hardware capabilities in supporting future voice services.
- Device service test: This test targets at access nodes that are already provisioned with the voice service. Before voice services are provisioned from the access node to a user, a device service test is performed to determine whether the voice service capabilities are supported by the access node.

Feature Dependency and Limitation

- Do not pick up the phone during a loop test. Otherwise, the test results will be incorrect.
- If the dialup mode of a PSTN port is set to **DTMF-only**, no loop test can be started on the PSTN port.

Device Hardware Test

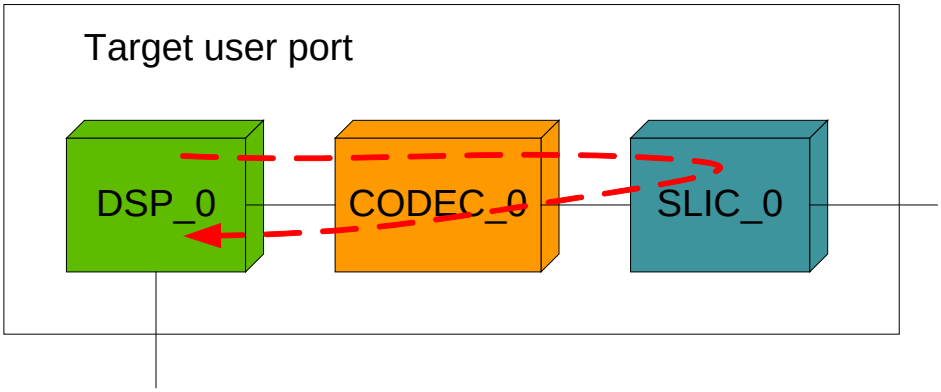
The device hardware test involves the following items:

- Off-hook detection
- On-hook detection
- Ringing and ringing stopping detection on a service port
- Speech path detection

For the first 3 test items, a POTS board emulates off-hook, on-hook, ringing, and ringing stopping, while the control board of the device performs the loop test on the POTS board. The last test item (speech path detection) is used to verify the service processing capability of the chip through service loopbacks. A speech path detection involves 3 loopback tests: SLIC loopback test, Codec loopback test, and DSP loopback test.

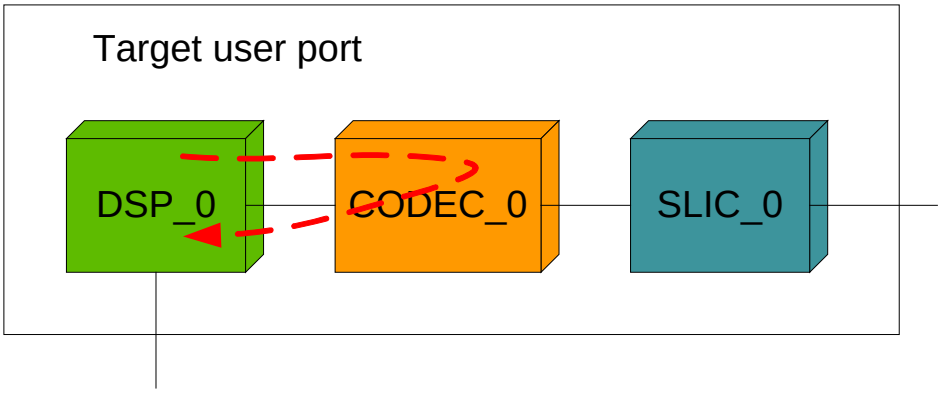
- In a SLIC loopback test, a POTS port is connected to a DSP channel, and the DSP chip generates the DTMF authentication code, which is sent to the TDM side. The DSP chip then detects the DTMF returned from the TDM side to check whether the channel between the DSP, Codec, and SLIC is normal, as shown in Figure 23-101.

Figure 23-101 Working principles of a SLIC loopback test



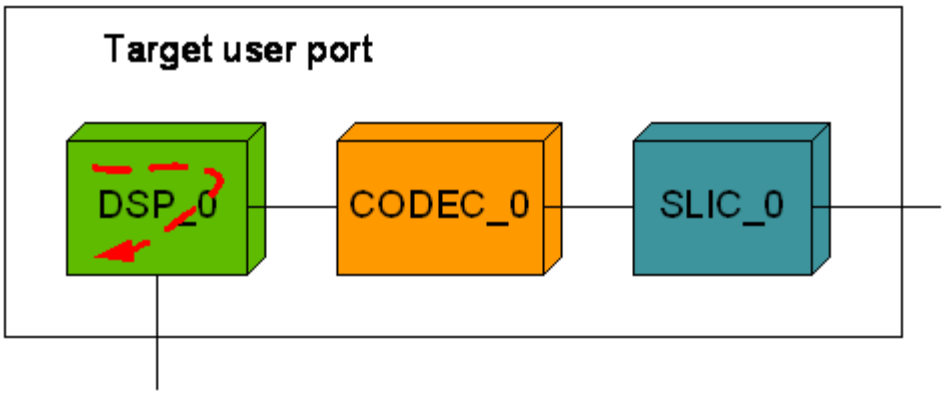
- In a Codec loopback test, the Codec loopback is set to the network side (remote loopback), and the DSP chip generates the DTMF authentication code, which is sent to the TDM side. The DSP chip then detects the DTMF returned from the TDM side to check whether the channel between the DSP and the Codec is normal, as shown in Figure 23-102.

Figure 23-102 Working principles of a Codec loopback test



- In a DSP loopback test on the TDM side, a DSP channel is looped back from the TDM side to the IP side, and the DSP chip generates the DTMF authentication code, which is sent to the TDM side. The DSP chip then detects the DTMF returned from the TDM side to verify the transmit and receive functions of the TDM side, as shown in Figure 23-103.

Figure 23-103 Working principles of a DSP loopback test on the TDM side



NOTE

In test mode, run the **pots path-test** *frameid/slotid* **port** *portid* **type hardware** command to start a device hardware test. The test results will be displayed after the test is completed. Engineers diagnose whether the system functions normally based on the test results.

Device Service Test

The principles of a device service test are similar to those of a call emulation test. In a device service test, the device acts as both the calling party and called party (that is, the dest port on the device calls another assistant port on the same device), and the system checks whether the configurations are correct. User ports do not enter the conversation state during a POTS port loop test. When the called port detects the ringing, the calling port emulates user on-hook, and therefore, no charge record is generated. A device service test checks:

- Device interface data, including the signaling data (such as MG interface data, protocol parameters, and call server parameters)
- User port data, including the media gateway (MG) IDs and terminal IDs (TIDs) of H.248 users and user accounts and service rights of SIP users

NOTE

In test mode, run the **pots path-test** *frameid/slotid* **port** *portid* **type business** command to start a device service test. The test results will be displayed after the test is completed. Engineers diagnose whether the system functions normally based on the test results. For example, if the test result "The telno of dest port 0/4/0 is error" is displayed, maintenance engineers need to check whether the user and telephone number configured for port 0/4/0 are correct.

23.15.5 Search Tone Test

A search tone test is a simple line fault locating function intended for maintenance engineers. In a search tone test, the test module sends voice signals with the specific frequency and amplitude to a line, and then maintenance engineers use a receiver or a dedicated device to detect the signals on the line. In addition, search tone tests can help maintenance engineers pinpoint the specific line among multiple user lines.

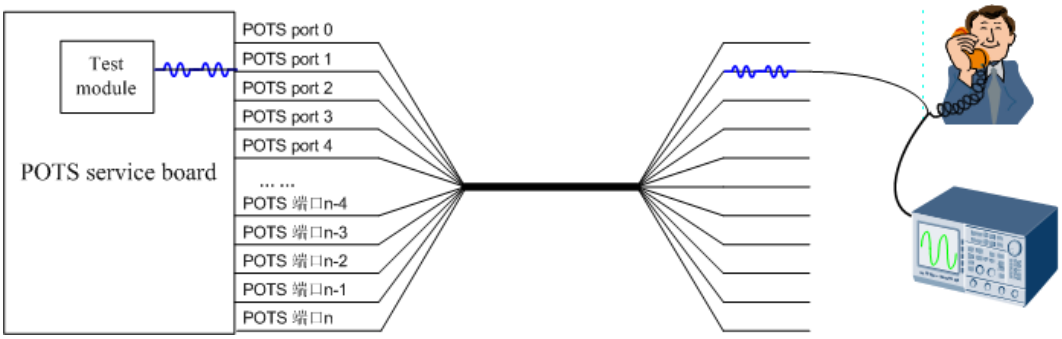
NOTE

- A search tone test can be performed even when a port is powered off.
- A search tone test cannot be performed when another test task (such as a POTS user circuit test, POTS user loop line test, call emulation test, signal tone test, or POTS port loop test) is in progress.
- A search tone test cannot be performed when a port is in the prohibit state.

Test Procedure

Figure 23-104 shows the principles of a search tone test.

Figure 23-104 Principles of a search tone test



1. The test module sends voice signals with the specific frequency and amplitude (as underlined by blue wave lines in Figure 23-104) to a line after a search tone test is started.

   In test mode, run the **pots search-tone-test** *frameid/slotid/portid* **test-flag enable** command to start a search tone test. If users are making calls during a search tone test, maintenance engineers can cancel or forcibly perform the test based on actual conditions.

NOTE

If maintenance engineers forcibly perform the test, services on the port are interrupted and users' telephone services are affected. Therefore, exercise caution when performing this operation.

2. Maintenance engineers use a receiver or a dedicated device connected to the other end of the line to detect whether the voice signals can be received. If the voice signals can be received, the line works normally.
3. (Optional) Stop a search tone test.

   Run the **pots search-tone-test** *frameid/slotid/portid* **test-flag disable** command to stop a search tone test.

NOTE

In a search tone test, the duration of playing the search tone needs to be set based on actual requirements.

- If the preset duration does not expire, maintenance engineers can run this command to manually stop the test.
- If the preset duration has expired, the system automatically stops the test.

23.15.6 Signal Tone Test

In a signal tone test, the system sends the signal tone signals to a specific port of a POTS board and makes the port loop back the signals, and then checks whether the loopback signals can be detected. This test function helps maintenance engineers check whether the system can normally process the detection of the user off-hook and signal tone and locate hardware faults related to the user off-hook and signal tone playing.

A signal tone test includes the following types, as listed in the following table.

Test Type	Requiring Coordination of Upper-Layer Device?

Test Type	Requiring Coordination of Upper-Layer Device?
Busy tone test	The coordination of the upper-layer device is not required.
Ringback tone test	
Dial tone test	The dial tone test and special dial tone test include the following modes. • **out-of-service**: In this mode, the test is performed on the device, which is independent of the upper-layer device. If the device is not connected to the upper-layer device, use this mode. • **in-service**: In this mode, the device coordinates with the upper-layer device to perform the test. The upper-layer device controls the playing of the dial tone or special dial tone.
Special dial tone test (that is, the dial tone test started by the device based on different service requirements, such as call forwarding-unconditional)	

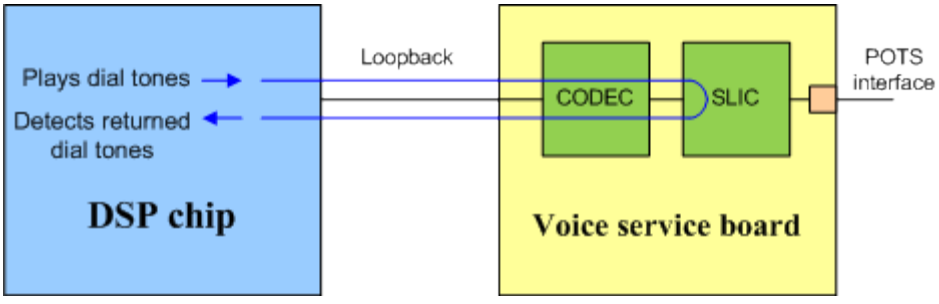
NOTE

- The POTS port supports local loopbacks.
- The POTS port must be in the on-hook state during a signal tone test.
- A signal tone test cannot be performed if the port is in the **prohibit** or **loopback** state.
- A signal tone test cannot be performed when another test task (such as a POTS user circuit test, POTS user loop line test, call emulation test, search tone test, or POTS port loop test) is in progress.

Test Procedure

Figure 23-105 shows the test procedure of a signal tone test. As shown in Figure 23-105, the subscriber line interface circuit (SLIC) supplies feeding current to the telephone, sends voice frequency, generates ringing, detects off-hook signals and on-hook signals, and processes analog signals. The CODEC performs the conversion between analog signals and digital signals.

Figure 23-105 Test procedure of a signal tone test



1. Maintenance engineers start a signal tone test on a POTS user port by using the CLI or NMS.
   - In test mode, run the **pots signal-tone-test** *frameid/slotid/portid* **signal-tone busy-tone** command to start a busy tone test.

- In test mode, run the **pots signal-tone-test** *frameid/slotid/portid* **signal-tone ringback-tone**command to start a ringback tone test.
- In test mode, run the **pots signal-tone-test** *frameid/slotid/portid* **signal-tone dial-tone**command to start a dial tone test.
- In test mode, run the **pots signal-tone-test** *frameid/slotid/portid* **signal-tone special-dial-tone** command to start a special dial tone test.

2. The DSP chip plays the signal tone.
   - For the busy tone test or ringback tone test, the device requests the DSP chip to issue the busy tone or ringback tone to the POTS port after the test is started.
   - For the dial tone test or special dial tone test:
     - If the test is started in **out-of-service** mode, the POTS port emulates the user off-hook. After the device detects the off-hook signal, it requests the DSP chip to issue the dial tone or special dial tone to the POTS port.
     - If the test is started in **in-service** mode, the POTS port emulates the user off-hook. After the device detects the off-hook signal, it reports the signal to the softswitch, and the softswitch requests the DSP chip to issue the dial tone or special dial tone to the POTS port.
3. The POTS board loops back the signal tone signals (as shown by the blue lines in Figure 23-105).
4. If the DSP chip can detect the loopback signals, the system runs normally.

   If any exception occurs during the test, the following methods can be used for troubleshooting.

Exception	Troubleshooting Method
No looped back signal is detected.	• Run the **display pstn state** command to query the port status to check whether the port is faulty. • Run the **display pstn state** command to query the port status to check whether the port is busy. • Run the **display dsp state** command to query the DSP channel status to check whether the DSP resources are sufficient.
The delay of the signal tone is too long.	• Run the **display cpu** command to query the CPU usage to check whether the system is overloaded. • If the test is started in **in-service** mode, check whether the interaction between the device and the softswitch is delayed.

5. (Optional) Stop a signal tone test.

   Run the **pots signal-tone-test** *frameid/slotid/portid* **test-flag disable** command to stop a signal tone test.

NOTE

In a signal tone test, the duration of playing the signal tone needs to be set based on actual requirements.

- If the preset duration does not expire, maintenance engineers can run this command to manually stop the test.
- If the preset duration has expired, the system automatically stops the test.

23.15.7 RTCP Statistics

Complying with the H.248 protocol, the softswitch, during and at the end of a call, can query the RTCP statistics of a user, including the number of transmitted RTP packets, bytes of transmitted RTP packets, number of received RTP packets, bytes of received RTP packets, number of lost transmitted packets, number of lost received packets, network jitter, and network loop delay.

The MA5600T/MA5603T/MA5608T reports its real-time statistics to the softswitch when the softswitch issues signaling to query the statistics. Then, the softswitch or the OSS system can manage the quality monitoring based on the statistics.

23.15.8 Signaling Tracing

This topic describes the purpose and implementation principles of signaling tracing fault diagnosis feature.

NOTE

Based on your requirements, VBD fault diagnosis may obtain some contents of users' communications (integrity communication contents are not obtained and user information will not be disclosed) for the purpose of safeguarding network operations and protecting services. Huawei alone is unable to collect or save the content of users' communications. You must comply with the laws and regulations of the countries concerned for using the VBD fault diagnosis feature. You are obligated to take considerable measures to ensure that the content of users' communications is fully protected when the content is being used and saved.

Introduction

After the OLT is running on the network and voice service failures occur, such as poor communication quality, low fax/modem service success rate, one-way audio, and no audio, easy signaling tracing (instead of remote packet capture) can be used to obtain voice signaling streams and further these signaling streams can be used for fast fault locating.

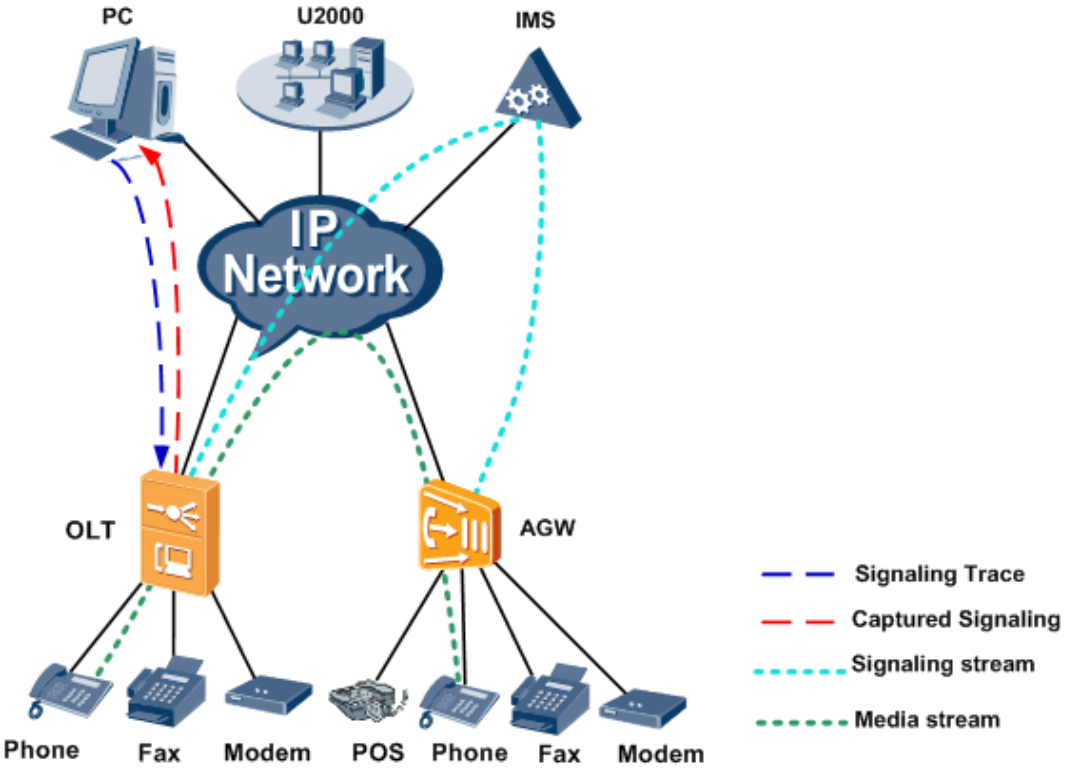
Signaling tracing features:

- Remote fault locating: Fault locating can be remotely implemented, eliminating the need of site visits.
- Fast fault locating: Remote management or packet capture software is not needed, shortening fault locating duration.
- Accurate fault locating: Voice signaling tracing can be performed based on fault symptoms.
- High security and reliability: The obtained signaling file does not contain any user privacy, such as the dialed telephone number, name, SMS, and FSK.

Networking Applications

Figure 23-106 shows the networking of signaling tracing.

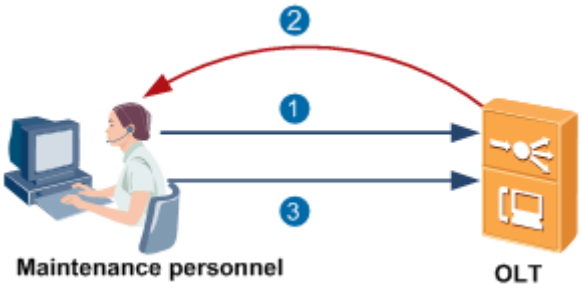
Figure 23-106 Networking of signaling tracing



Signaling tracing takes effect on signaling exchanged between the IMS and OLT. The O&M engineer obtains the traced signaling by telnetting to the OLT from the PC.

Figure 23-107 shows the process of obtaining voice signaling.

Figure 23-107 Process of obtaining voice signaling

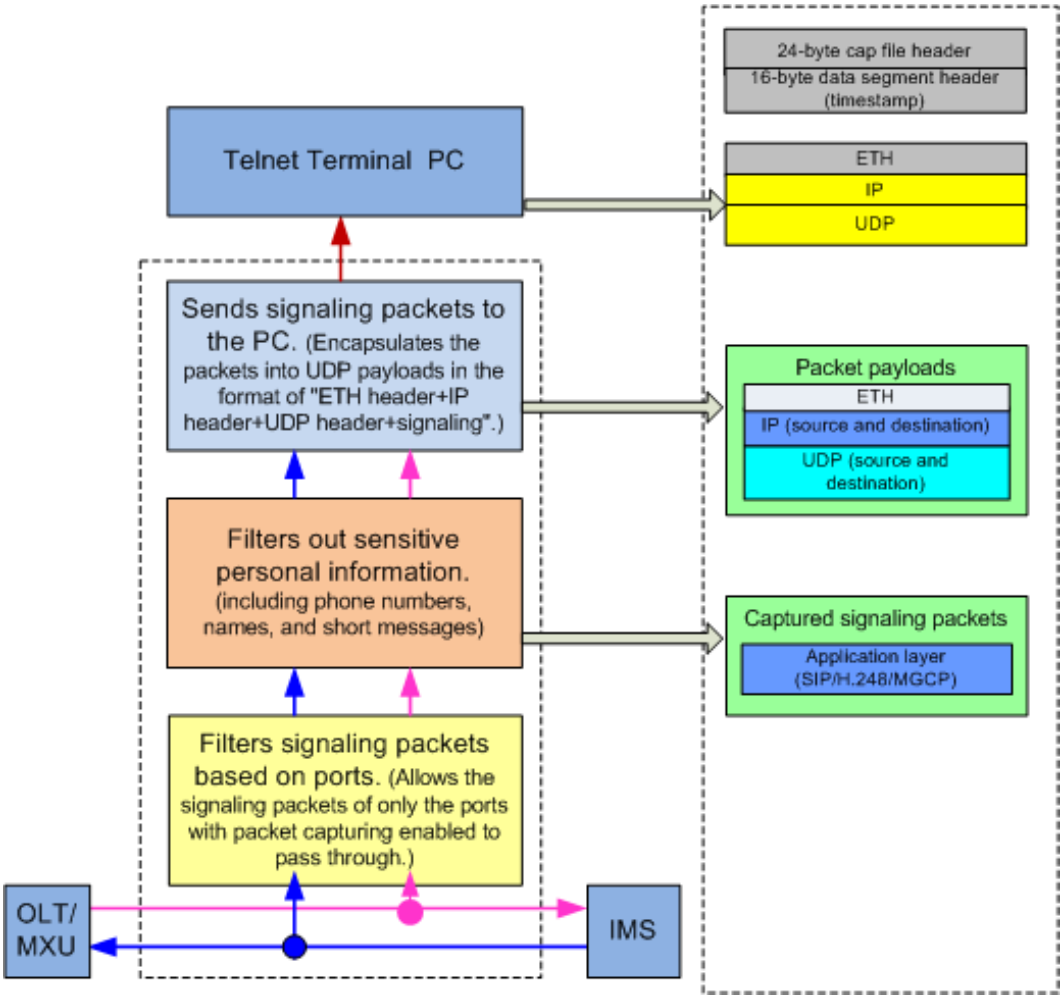


1. The O&M engineer telnets to the OLT from the PC and starts voice signaling tracing through the CLI.
2. The OLT sends the obtained voice signaling to the specified PC.
3. The O&M engineer stops obtaining voice signaling.

Working Principles

Figure 23-108 shows the application of authorized signaling tracing.

Figure 23-108 Application of authorized signaling tracing



After authorized signaling tracing is enabled:

1. The OLT filters packets based on ports. The OLT captures the signaling packets only on the ports with signaling packet capturing enabled.
2. The OLT filters out sensitive personal information, including phone numbers, names, short messages, FSK data, and DTMF data, from the captured signaling packets.
3. The OLT encapsulates the signaling packets into UDP payloads in the format of "ETH header+IP header+UDP header+signaling" and sends the UDP payloads to the PC.
4. The O&M engineer analyzes the obtained voice signaling and locates faults on the telnet terminal.

Procedure

Perform the following to obtain the voice signaling to locate a voice service fault quickly.

1. Run the **diagnose** command to enter diagnose mode from privilege mode or global config mode.
2. Run the **signal trace** command to start a signaling tracing task.

3. Run the **undo signal trace** command stop signaling tracing manually or the system stops signaling tracing automatically (after the preset signaling tracing duration expires).

Result

After signaling tracing, run the **display signal trace** command to query the signaling tracing information about the current port.

23.15.9 VBD Fault Diagnosis

This topic describes the purpose and implementation principles of the voice band data (VBD) fault diagnosis feature.

NOTE

Based on your requirements, VBD fault diagnosis may obtain some contents of users' communications (the information cannot restore the integrity communications contents) for the purpose of safeguarding network operations and protecting services. Huawei alone is unable to collect or save the content of users' communications. You must comply with the laws and regulations of the countries concerned for using the VBD fault diagnosis feature. You are obligated to take considerable measures to ensure that the content of users' communications is fully protected when the content is being used and saved.

Introduction

The VBD fault diagnosis feature enables the OLT to obtain voice packets, signaling packets, and voice control packets when a narrowband service fault occurs on the OLT. Then, O&M engineers can use these packets for rapid fault location. The narrowband service fault can be:

- Poor quality of voice communication
- High failure ratio of the fax or modem service
- One-way audio or no audio

NOTE

Voice control packets contain higher-layer communication protocols used between the control board and service boards of the OLT as well as digital signal processor (DSP) commands for processing internal board data.

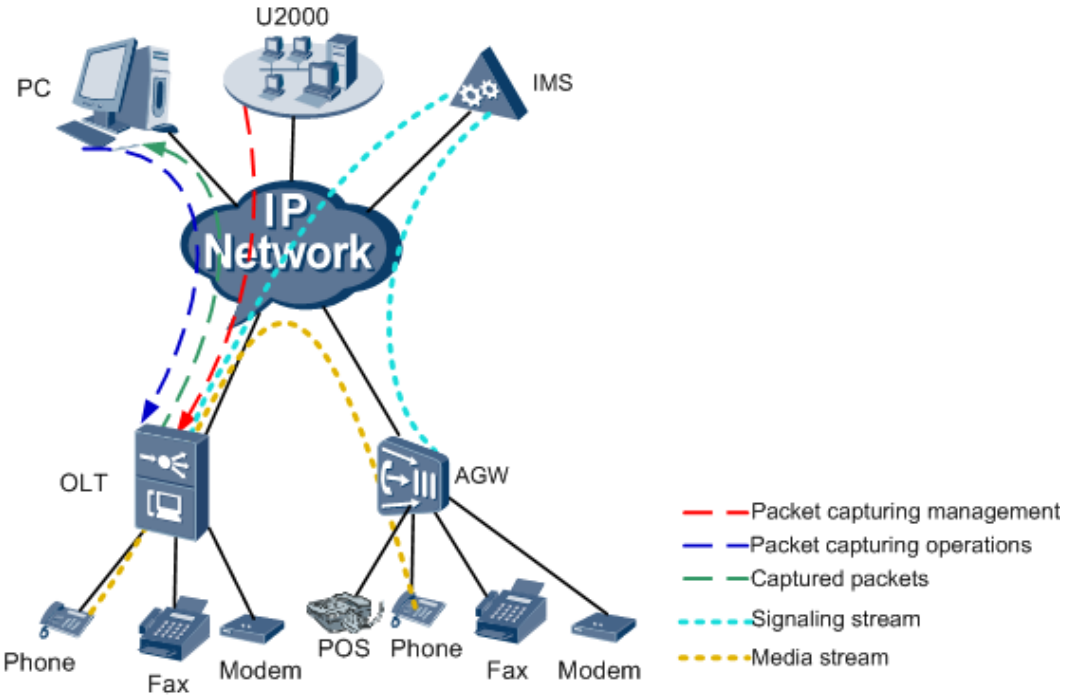
The VBD fault diagnosis feature supports:

- Remote fault location: removes the need of onsite fault location.
- Rapid fault location: enables O&M personnel to obtain voice packets, signaling packets, and voice control packets based on fault symptoms, thereby improving fault location efficiency.
- Secure and reliable data obtaining: The obtained voice packets, signaling packets, and voice control packets do not contain any user privacy information, such as called numbers, names, short messages, or frequency shift keying (FSK) data.

Networking Applications

Figure 23-109 shows the networking of the voice VBD fault diagnosis feature.

Figure 23-109 Networking of the VBD fault diagnosis feature



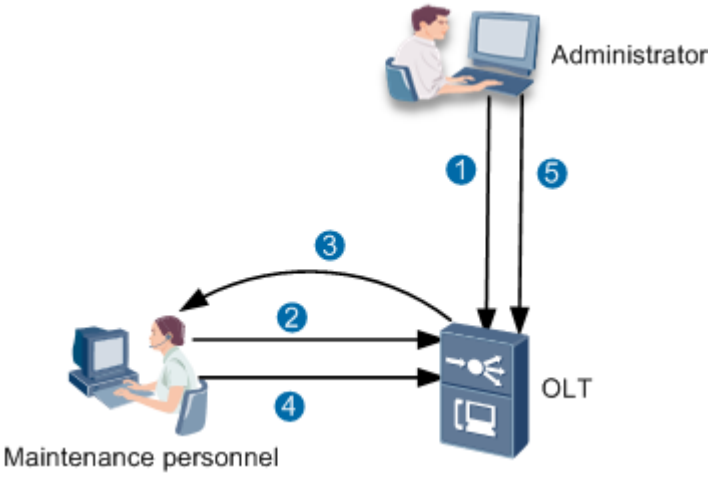
Carrier's administrator grants packet capturing rights to the O&M personnel. Then, the O&M personnel capture packets on a specified PC (packet capturing server). The OLT filters out users' communication content from the captured packets, encapsulates the filtered packets into UDP packet payloads, and sends the UDP packets to the PC.

NOTE

Authorized packet capturing complies with local laws and regulations and is performed carrier's authorization. Authorized packet capturing is used to capture voice packets, signaling packets, and voice control packets that do no contain user sensitive information and the captured data is only used for fault location.

Figure 23-110 shows the process of obtaining voice packets.

Figure 23-110 Process of obtaining voice packets



1. The carrier's administrator uses the administrator account to remotely log in to the OLT from the PC and grants packet capturing rights to the O&M personnel.
2. The O&M personnel enable a general voice packet capturing tool, such as wireshark, on the PC, use a non-administrator account to remotely log in to the OLT, capture voice packets through the maintenance port.

NOTE
To safeguard network operations and protect services, the O&M personnel can use only a non-administrator account to log in to the OLT.

3. The OLT sends the obtained voice packets to the PC.
4. The O&M personnel stop capturing voice packets.
5. The carrier's administrator cancels the packet capturing permission.

Basic Concepts Related to Authorized Media Stream Packet Capturing

Capturing all voice packets of a call easily discloses user privacy information. To protect user privacy, the OLT provides differentiated VBD fault diagnosis modes. By doing so, VBD fault diagnosis can be enabled without bringing the risk of disclosing user privacy information.

Table 23-23 lists the VBD fault diagnosis modes that the OLT provides.

Table 23-23 VBD fault diagnosis modes

Mode	Description	Usage Scenario
Full packet capturing	The OLT does not filter or replace any information in the captured Real-Time Transport Protocol (RTP) or time division multiplexing (TDM) tracing packets.	This mode applies when the captured packets do not contain the content of users' communications. In this mode, packets can be captured as many as possible, which facilitates fault location. This mode is used to diagnose fax or modem negotiation failures.
Fuzzy packet capturing	The OLT retains the minimum data that identifies the fax or modem signal tone in RTP and	This mode applies when the OLT cannot determine whether a user will initiate a VBD call, or when the content

Mode	Description	Usage Scenario
	TDM tracing packets, which prevents the original content of users' communications from being restored. Specifically, the OLT retains only 10 ms of data (data generated in a 10-millisecond duration) from every 80 ms of data and erases the 70-ms data. In this way, packets are captured in a discontinuous way and the original content of users' communications cannot be restored.	of users' communications will be transmitted after a VBD call is set up. This mode is used when the local or peer end cannot identify a fax or modem signal tone or a fax or modem negotiation between the local and peer ends fails.
Packet header-only capturing	The OLT captures only the IP, UDP, and RTP headers of RTP packets. It replaces the payloads of the RTP packets with fixed data. In addition, the OLT does not capture TDM tracing packets.	This mode applies when the VBD negotiation process ends. This ensures that no content of users' communications will be captured. This mode is used to locate voice quality faults, such as packet loss, jitter, delay, one-way audio, and no audio.

Working Principles of Authorized Media Stream Packet Capturing

Figure 23-111 and Figure 23-112 show the application of the voice packet capturing modes in a fax or modem call and in a common call (Taking the local ringback tone as an example).

Figure 23-111 Application of the voice packet capturing modes in a fax or modem call

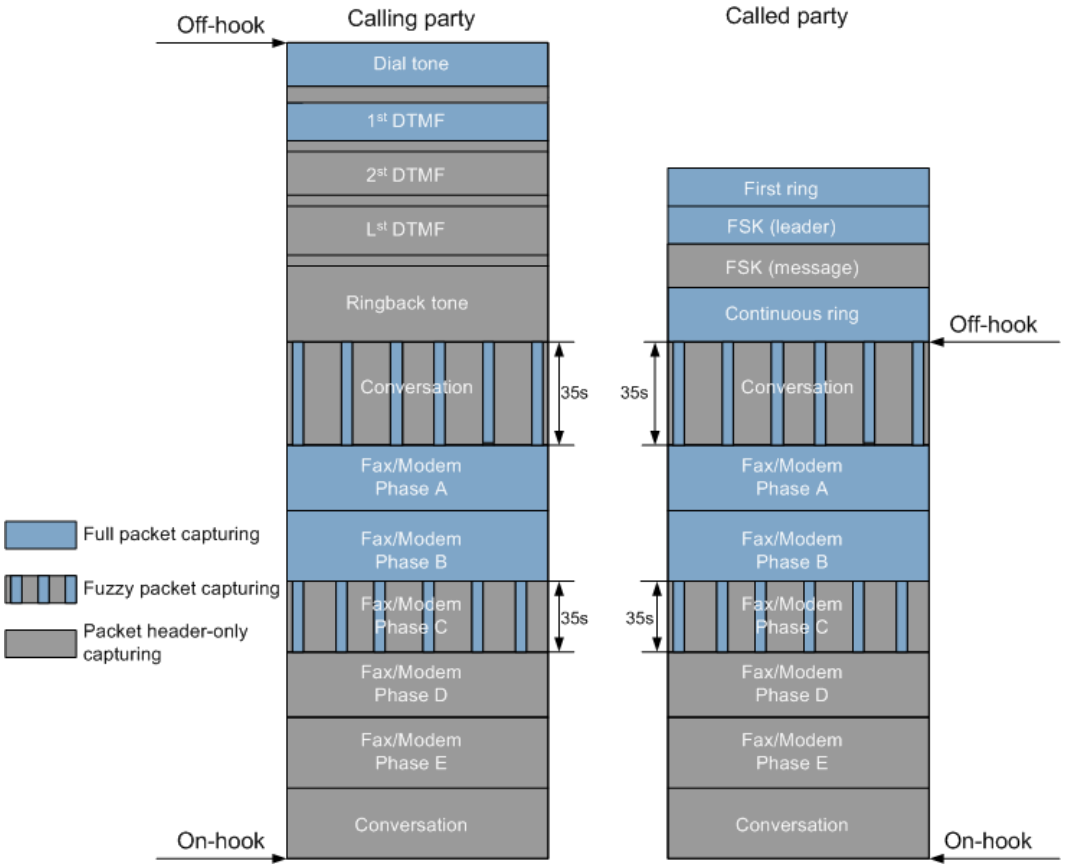
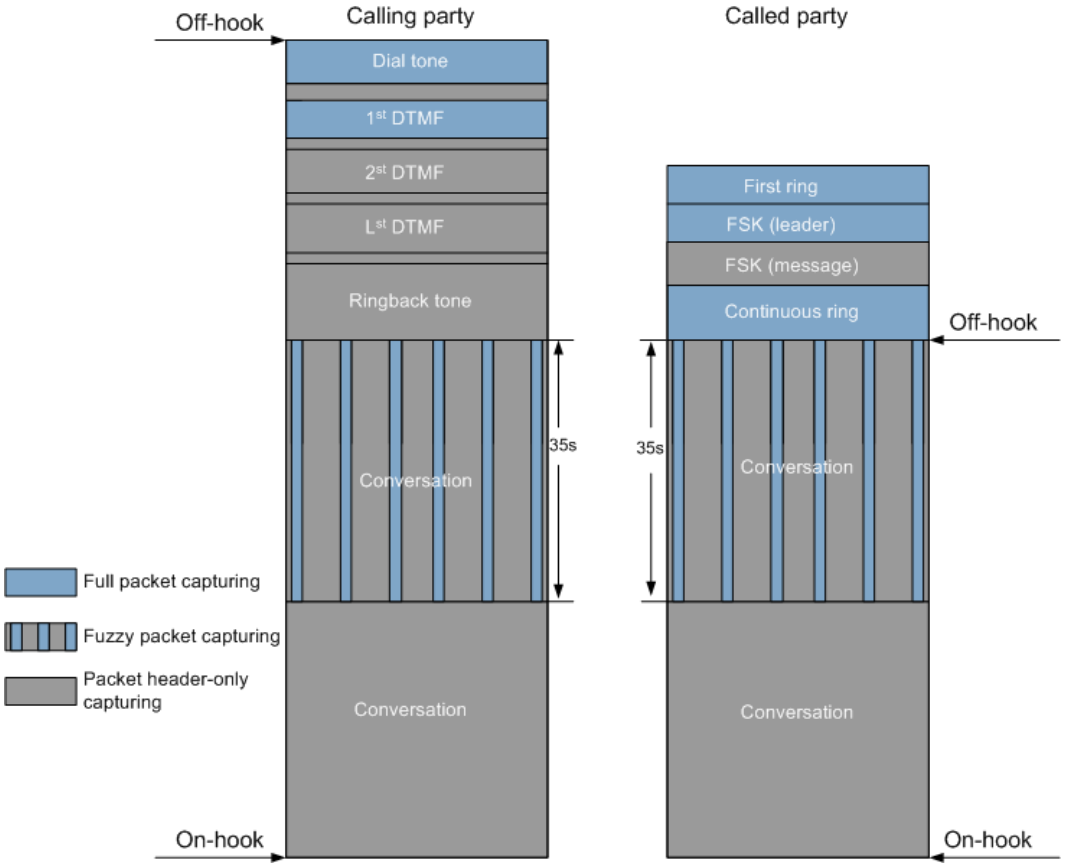


Figure 23-112 Application of the voice packet capturing modes in a common call



For the calling party:

1. After the calling party picks up the phone, the OLT starts the full packet capturing mode.
2. After detecting the first dual tone multiple frequency (DTMF) number, the OLT switches to the packet header-only capturing mode, which prevents disclosure of the user DTMF numbers.
3. When the calling party hears ringback tones, the capturing mode does not changed.
4. After the call is set up between the calling and called parties, the OLT switches to the fuzzy packet capturing mode and starts a fuzzy packet capturing protection timer (35s). The fuzzy packet capturing protection timer limits the fuzzy packet capturing duration and protects user communication security.

NOTE

The call may be a VBD call, during which the VBD signal tone may not be identified. Therefore, the fuzzy packet capturing mode is required to locate faults where the VBD signal tone fails to be identified.

5. If the OLT detects the fax or modem signal tone and the two ends of the fax or modem service are at the negotiation phase, the OLT switches to the full capturing mode because no communication data is involved at the negotiation phase. In this way, the OLT can capture most fault-related packets, which facilitates rapid fault location.

NOTE

This step is involved only in a fax or modem call.

6. Before the negotiation phase ends, the OLT switches to the fuzzy packet capturing mode.

NOTE
This step is involved only in a fax or modem call.

7. When the fuzzy packet capturing protection timer times out, the OLT switches to the packet header-only capturing mode, which protects user communication security.

For the called party:

1. When the phone of the called party rings, the OLT starts the full packet capturing mode.
2. When sending FSK data to the called party, the OLT switches to the packet header-only capturing mode, which prevents the content of users' communications from being captured.
3. After the call is set up between the calling and called parties, the OLT switches to the fuzzy packet capturing mode and starts a fuzzy packet capturing protection timer (35s). The fuzzy packet capturing protection timer limits the fuzzy packet capturing duration and protects user communication security.

NOTE
The call may be a VBD call, during which the VBD signal tone may not be identified. Therefore, the fuzzy packet capturing mode is required to locate faults where the VBD signal tone fails to be identified.

4. If the OLT detects the fax or modem signal tone and the two ends of the fax or modem service are at the negotiation phase, the OLT switches to the full capturing mode because no communication data is involved at the negotiation phase. In this way, the OLT can capture most fault-related packets, which facilitates rapid fault location.

NOTE
This step is involved only in a fax or modem call.

5. Before the negotiation phase ends, the OLT switches to the fuzzy packet capturing mode.

NOTE
This step is involved only in a fax or modem call.

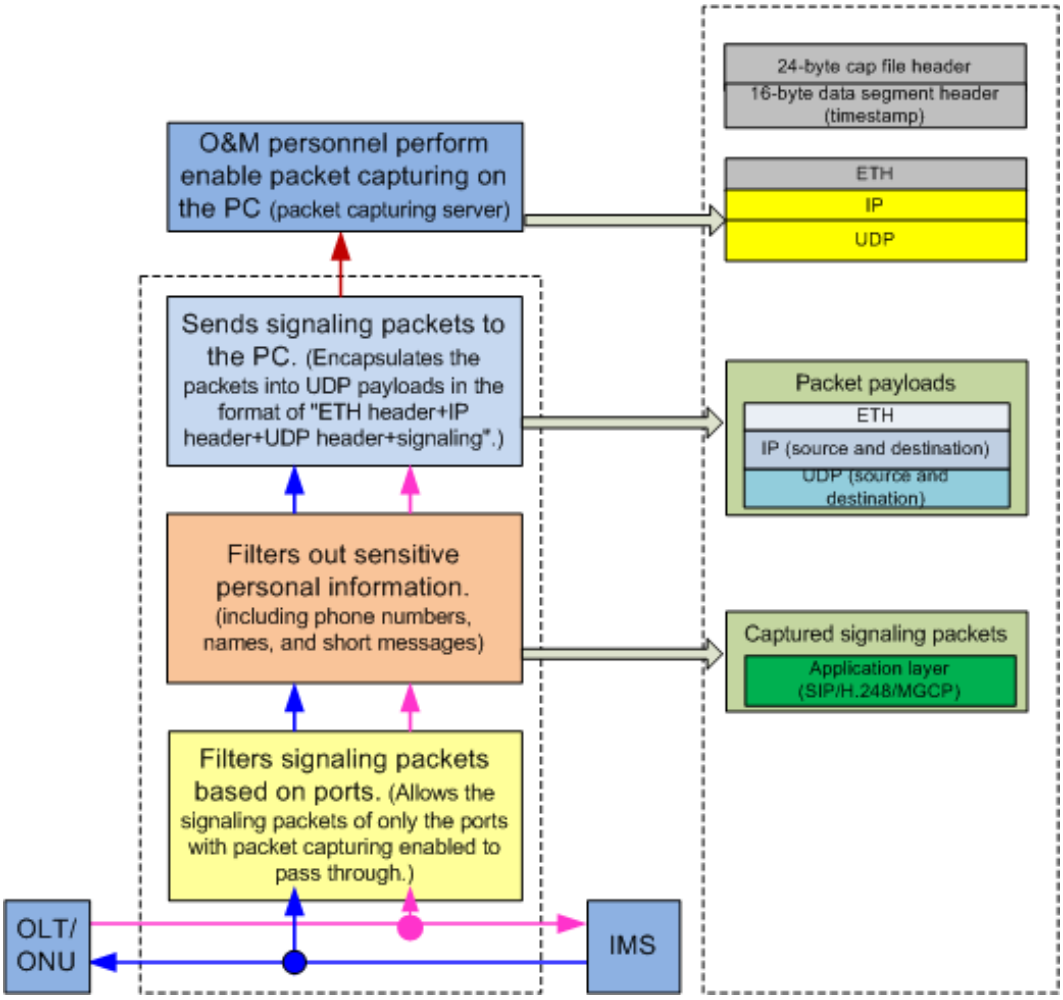
6. When the fuzzy packet capturing protection timer times out, the OLT switches to the packet header-only capturing mode, which protects user communication security.

NOTE
When detecting the first DTMF number (or an RFC 2833-compliant DTMF number), the OLT switches to the packet header-only capturing mode to protect the numbers dialed by users.

Working Principles of Authorized Signaling Packet Capturing

Figure 23-113 shows the application of authorized signaling packet capturing.

Figure 23-113 Application of authorized signaling packet capturing



After authorized signaling packet capturing is enabled, the OLT performs the following operations on the signaling packets, such as SIP, MGCP, or H.248 packets, received from the IP multimedia subsystem (IMS) or received by the OLT:

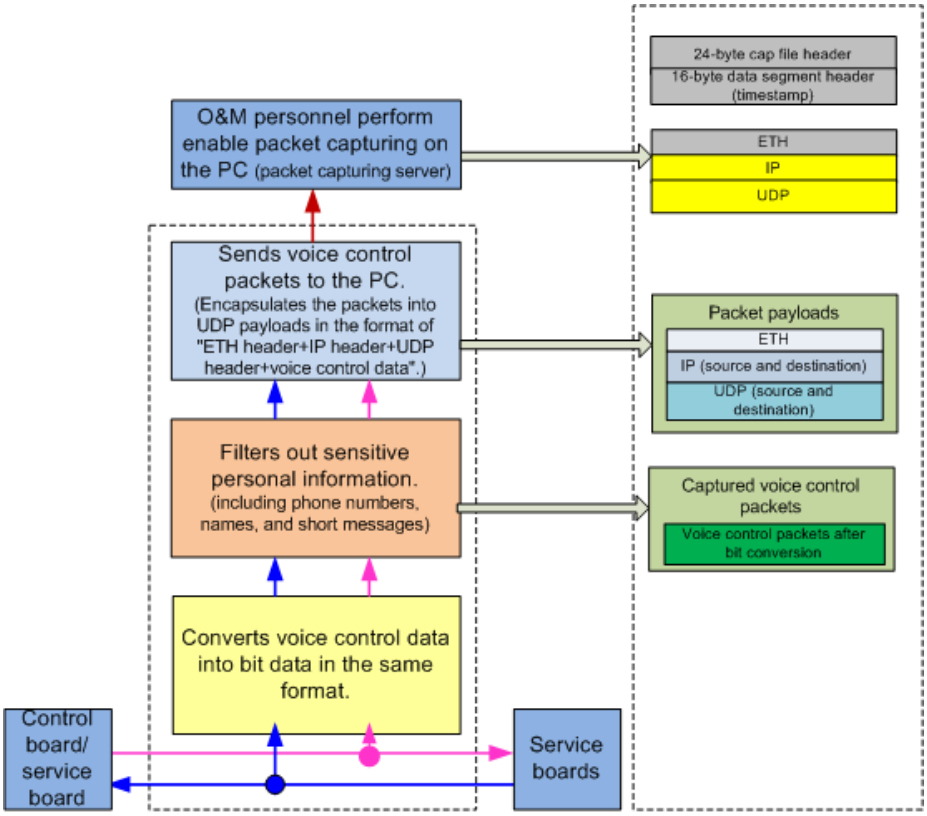
1. Filters packets based on ports. The OLT captures the signaling packets only on the ports with signaling packet capturing enabled.
2. Filters out sensitive personal information, including phone numbers, names, short messages, FSK data, and DTMF data, from the captured signaling packets.
3. Encapsulates the signaling packets into UDP payloads in the format of "ETH header+IP header+UDP header+signaling" and sends the UDP payloads to the PC.

The O&M personnel enable a general voice packet capturing tool, such as wireshark, on the PC to capture voice packets and analyze the voice packets for fault location.

Working Principles of Authorized Voice Control Packet Capturing

Figure 23-114 shows the application of authorized voice control packet capturing.

Figure 23-114 Application of authorized voice control packet capturing



After authorized voice control packet capturing is enabled, OLT performs the following operations on the packets transmitted between the control board and service boards or between service boards:

1. Converts voice control data into bit data in the same format.
2. Filters out sensitive personal information, including phone numbers, names, short messages, FSK data, and DTMF data, from the captured voice control packets.
3. Encapsulates the voice control packets into UDP payloads in the format of "ETH header+IP header+UDP header+voice control data" and sends the UDP payloads to the PC.

The O&M personnel enable a general voice packet capturing tool, such as wireshark, on the PC to capture voice packets and analyze the voice packets for fault location.

Configuration Procedure

Perform the following steps to capture voice packets:

1. Run the **remote-capture whitelist add** command to add a specified user port on the PC to the remote packet capturing whitelist.

   Only the ports in the remote packet capturing whitelist can be used to capture voice packets.

NOTE

The carrier's administrator must grant the operation rights to the O&M personnel for this step.

2. Run the **remote-capture parameters** command to set the IP address and port number of the PC where voice packets can be captured.
3. Run the **remote-capture trace** command to configure the port or channel for capturing packets and the packet capturing duration, and start voice packet capturing.
4. Run the **undo remote-capture trace** command to manually stop capturing voice packets.
5. After capturing voice packets, run the **remote-capture whitelist delete** command to delete the user port from the remote packet capturing whitelist.

Configuration Results

After the configuration, you can:

- Run the **display remote-capture state** command to query packet capturing status on the port.
- Capture voice packets on the PC where the general voice packet capturing tool, such as wireshark, is enabled.

23.15.10 QoS Alarm

The network quality affects the voice service to a great extent. During a call, the MA5600T/MA5603T/MA5608T monitors the network quality in real time. When the network quality is below the pre-set threshold, a corresponding alarm is generated on the MA5600T/MA5603T/MA5608T to warn the customer of the network quality.

The QoS alarm function is used to monitor three indexes, packet loss, loop delay, and jitter. The corresponding values can be set according to the actual network condition. During a call, the MA5600T/MA5603T/MA5608T collects the data of packet loss, loop delay, and jitter, and then compares the data with the preset thresholds. When the data exceeds the thresholds, an alarm is generated. When the network indexes return to be lower than the preset thresholds, a recovery alarm is generated on the MA5600T/MA5603T/MA5608T.

The QoS alarms can detect the network abnormalities in real time. When users complain, the QoS alarm can be referred to locate the fault (whether caused by the network or device).

23.16 Voice Reliability

This topic describes features related to voice reliability, including dual-homing networking, highly reliable transmission (SCTP), and voice QoS.

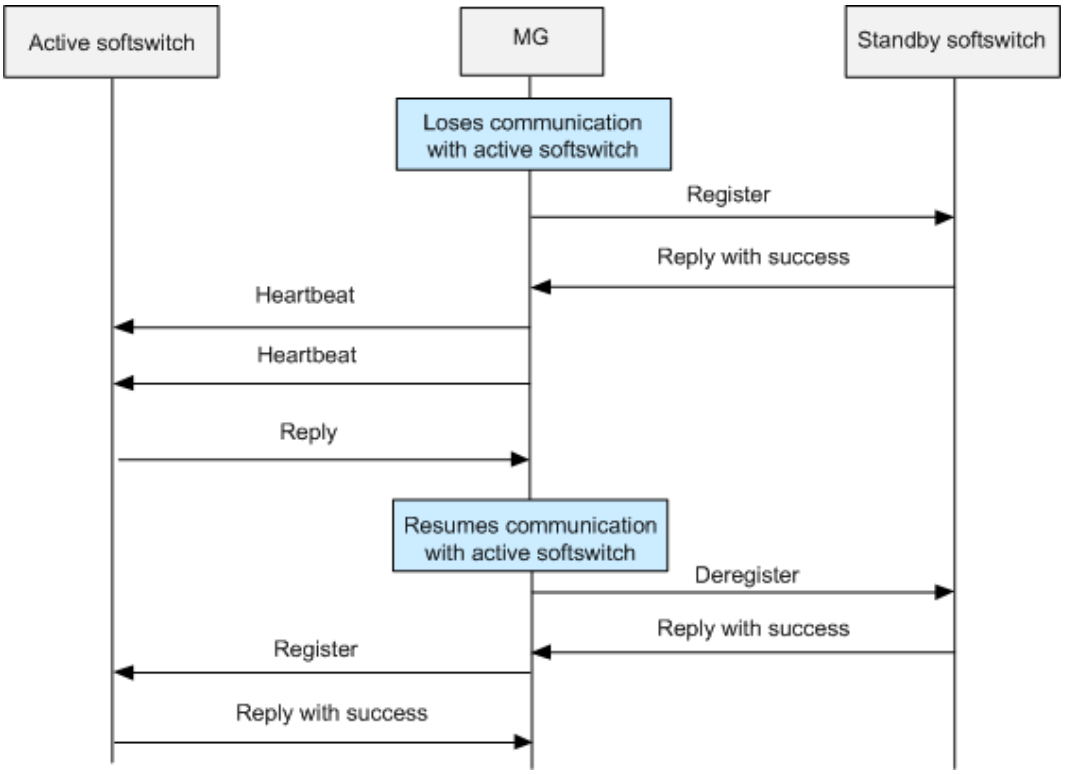
23.16.1 H.248/MGCP Dual Homing

Dual homing is an NGN (Next Generation Network) total solution. Based on this solution, when the active softswitch or the link from the MG to the active softswitch is faulty, the MG need be switched to the standby softswitch immediately to prevent call services of users connected to the softswitch and the MG from being affected.

Dual homing requires that one MG is configured with two softswitches, one active and one standby. The connection between the MG and the softswitch is detected through the heartbeat message.

Figure 23-115 illustrates the working principle of dual homing.

Figure 23-115 Working principle of dual homing



1. The MG detects the interrupted connection between the MG and the active softswitch through the heartbeat message.
2. The MG registers with the standby softswitch.
3. The MG sends the detection messages to the active softswitch at regular intervals (same as common heartbeat intervals), If the MG receives the response from the active softswitch, it indicates that the communication with the active softswitch is recovered. In this case, the MG takes the next action. If receiving no response from the softswitch, the MG keeps sending the detection messages.
4. The MG sends a message to the standby softswitch for service cancellation and waits for the response from the softswitch.
5. After receiving the response from the standby softswitch, the MG starts to register with the active softswitch. If three consecutive attempts of registration fail, the MG registers with the standby softswitch again following the same procedure.

Different carriers may choose the following different dual homing policies:

1. When the original active softswitch recovers, the MG automatically switches to the original active softswitch.
2. The MG does not support the auto-switching. Regardless of whether the MG registers with the active softswitch or the standby softswitch, if the softswitch with which the MG registers is normal, the MG works with this softswitch all along. The MA5600T/MA5603T/MA5608T can support the preceding two policies through related configuration. By default, the MA5600T/MA5603T/MA5608T supports the second policy.

23.16.2 H.248 Multi-homing

Overview

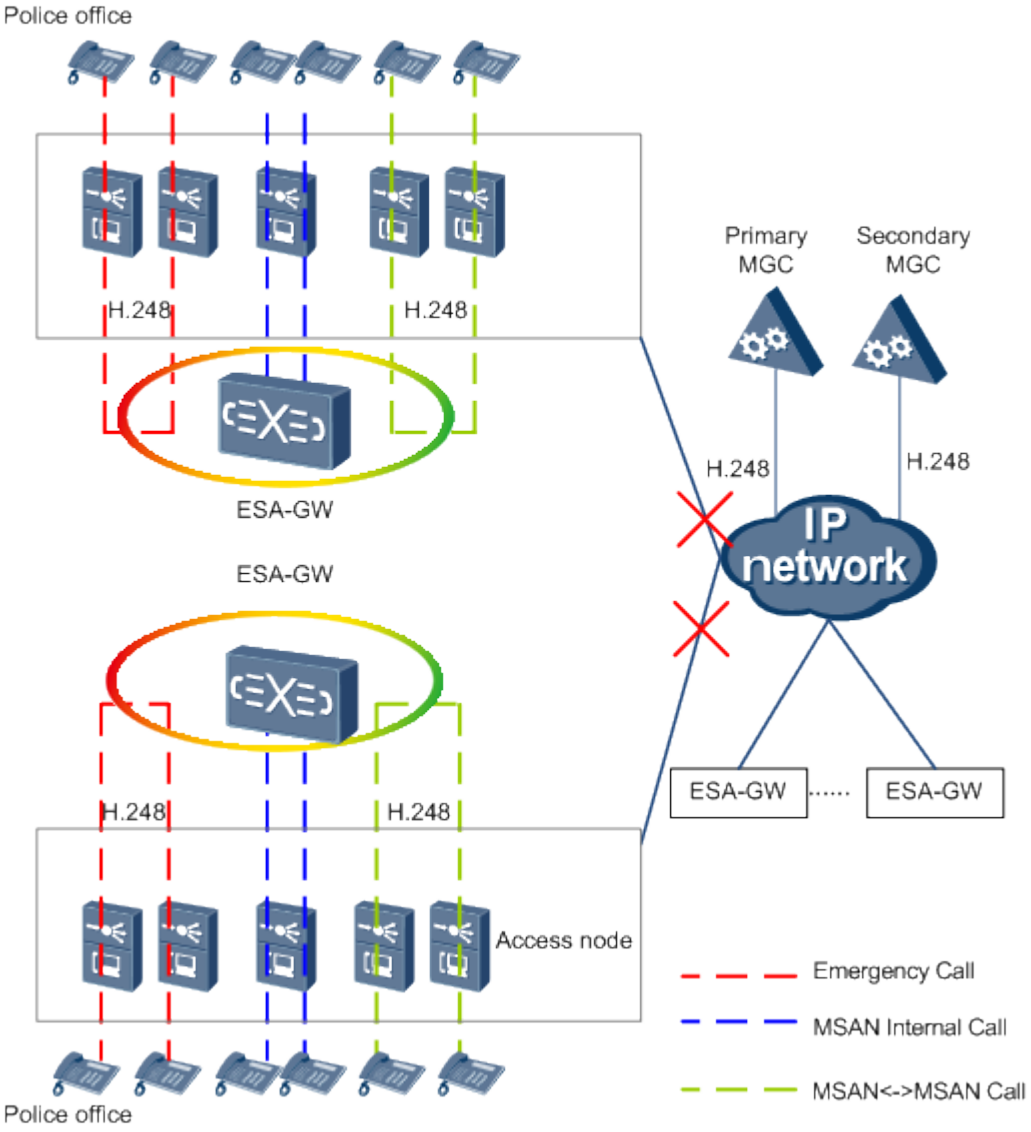
As an enhancement of dual-homing, multi-homing is a configuration in which a media gateway (MG) is homed to the primary media gateway controller (MGC), secondary MGC, and disaster-recovery MGC.

The system supports the following configurable switching policies for multi-homing:

1. Automatic switching back
   - An MG registering with the secondary MGC will automatically switch back to the primary MGC when the primary MGC recovers.
   - An MG registering with the disaster-recovery MGC will automatically switch back to the primary/secondary MGC when the primary/secondary recovers.
2. No automatic switching back
   - An MG registering with the secondary MGC will not automatically switch back to the primary MGC when the primary MGC recovers.
   - An MG registering with the disaster-recovery MGC will not automatically switch back to the primary/secondary MGC when the primary/secondary MGC recovers.

Network Application

Figure 23-116 H.248 multi-homing network



- ESA-GW: emergency standalone-gateway
- MSAN: multi-service access node

As shown in the preceding figure, an MSAN is an MA5600T/MA5603T/MA5608T and an ESA-GW can be a small-capacity softswitch in network deployment. Generally, ESA-GWs and MA5600T/MA5603T/MA5608Ts are deployed in the same telecommunications room. When disconnected from the primary and secondary MGCs (for example, due to a fiber cut), the MA5600T/MA5603T/MA5608T initiates registration to the ESA-GW. After the successful registration, all call services of the MA5600T/MA5603T/MA5608T are controlled by the ESA-GW. In this way, the following services are still available even if the MA5600T/MA5603T/MA5608T is disconnected from the primary and secondary MGCs:

1. Call services of users connected to the same MA5600T/MA5603T/MA5608T

2. Call services of users connected to different MA5600T/MA5603T/MA5608Ts (homed to the same ESA-GW)
3. Emergency call services

Pay attention to the following aspects when applying H.248 multi-homing:

1. The call service capabilities are restricted by the ESA-GW.
2. The callee of an emergency call (for example, police emergency call) must be connected to an MA5600T/MA5603T/MA5608T.

NOTE

This limitation is a supplement to the solution shown in Figure 23-116. Whether such an limitation takes effect depends on the actual core network topology.

3. Only the POTS users are supported (the ISDN users and other users are not supported).

Switching Process

Figure 23-117 Process of switching to the ESA-GW

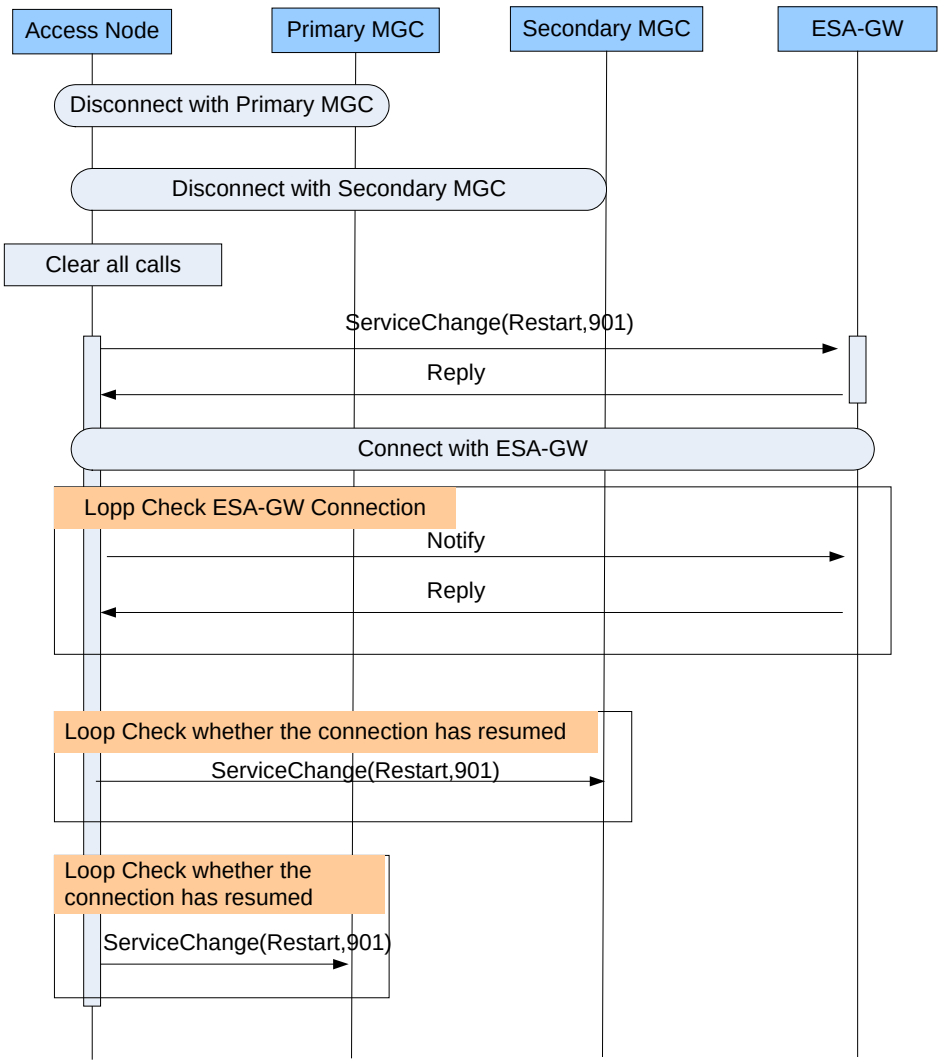
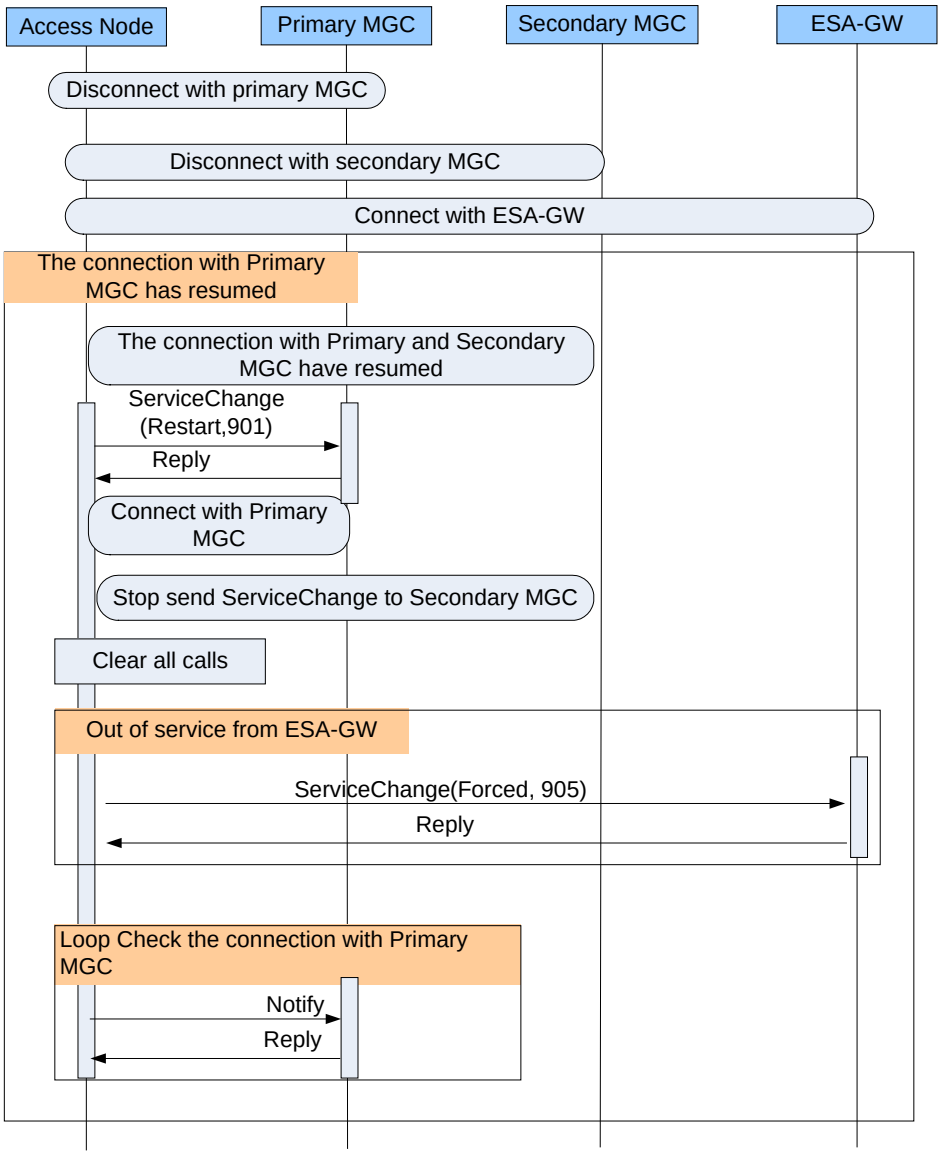


Figure 23-118 Process of switching back



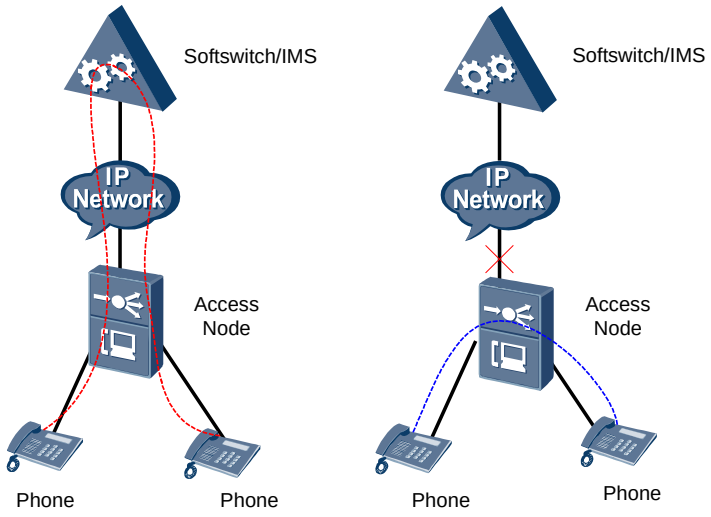
23.16.3 Emergency Standalone

Emergency standalone is a solution in which the users on the same MG can call each other even when the interface between the MG and the softswitch is interrupted. After a user picks up the telephone, the MG (namely, the MA5600T/MA5603T/MA5608T) checks whether the interface connected the softswitch is interrupted.

- If the interface is in the normal state, the normal softswitch process starts.
- Otherwise, the MG checks whether emergency standalone can be enabled.
  - If yes, the MA5600T/MA5603T/MA5608T controls the call process.
  - If no, the user listens to the busy sound (because the interface is faulty and emergency standalone is not allowed).

Figure 23-119 shows the operating principle of emergency standalone.

Figure 23-119 Operating principle of emergency standalone



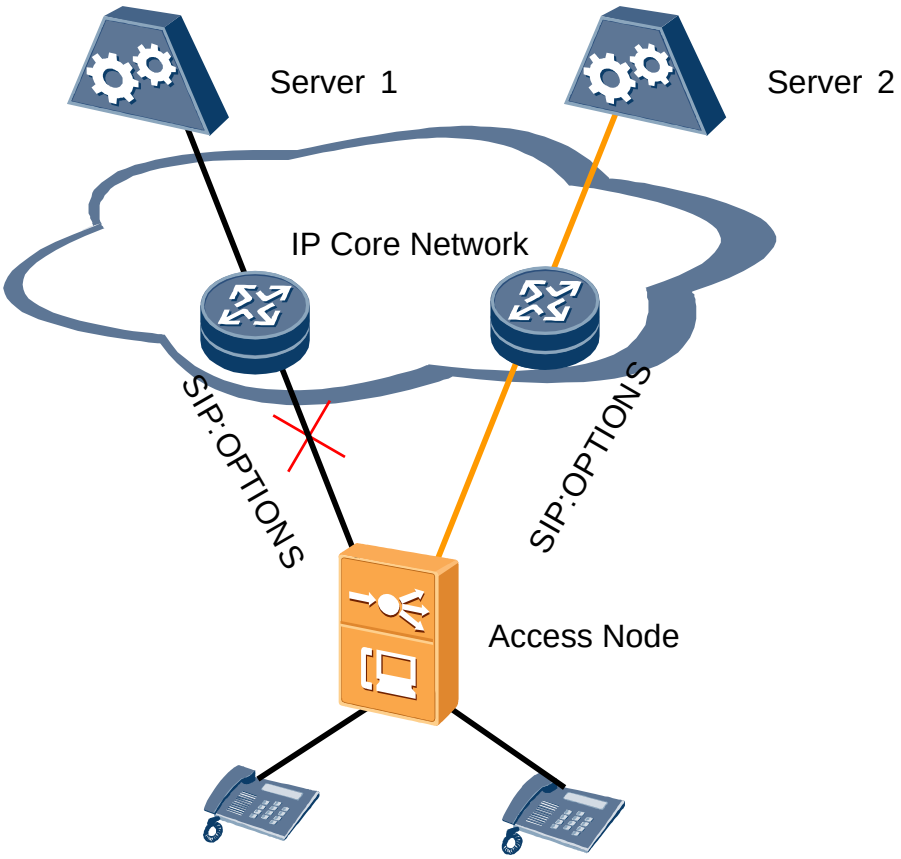
The emergency standalone process is as follows:

- The service processing after the calling party picks up the telephone is as follows:
  a. The calling party picks up the telephone.
  b. The device automatically delivers the dial tone to the calling party.
  c. The calling party dials a phone number.
  d. The device analyzes the dialed number and finds out the called party on the same device. The phone number is configured for a user when the user is configured on the MG.
  e. The device delivers the ringing signal and calling party's phone number to the called party.
  f. The device delivers the ring-back tone to the calling party.
- The service processing after the called party picks up the telephone is as follows:
  a. The called party picks up the telephone.
  b. The device stops delivering the ring-back tone to the calling party.
  c. The calling and called parties start a conversation.
- The service processing after any party puts down the telephone is as follows:
  a. Any of the two parties puts down the telephone.
  b. The device delivers the busy tone to the other party.
  c. The other party puts down the telephone.
- Limitation of emergency standalone:
  - Only the unabbreviated number is supported and the Centrex group, abbreviated number message, user outgoing/incoming call authority, and various new services are not supported.
  - A user can call another user only on the same VAG.
  - The feature applies only to the VoIP user.
  - The dual-homing feature and the emergency standalone feature cannot be enabled at the same time.

23.16.4 SIP Dual Homing

Figure 23-120 shows the networking of SIP dual homing.

Figure 23-120 Call releasing flow



The working flow of SIP dual homing is similar to the working flow of H.248/MGCP dual homing. The MA5600T/MA5603T/MA5608T detects the proxy server in real time. When the primary proxy server is faulty, services can be switched to the secondary proxy server. Before the switching, the call can be released. After the switching, the call can be initiated.

23.16.5 H.248/SIP over SCTP

Currently, most devices adopt H.248/SIP over UDP. H.248.4 recommends H.248/SIP over SCTP, which implements the message retransmission at the application layer through SCTP.

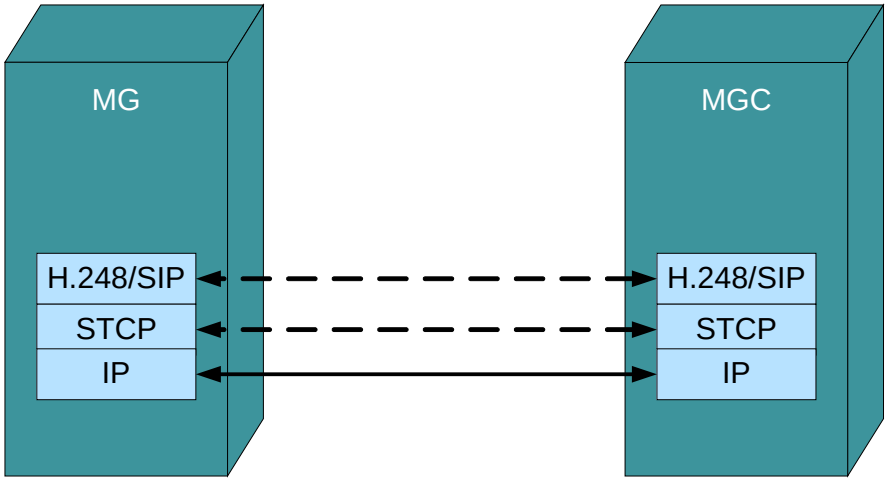
Compared with the UDP protocol, the SCTP protocol has the following advantages:

1. Reliability: Messages can be transmitted fast and reliably through SCTP.
2. Multi-homing: With the multi-homing feature, multiple IP addresses are supported on an SCTP endpoint. That is, an SCTP endpoint can use multiple physical network ports to enhance the endpoint reliability.
3. Congestion control: The congestion control through SCTP is similar to the congestion control through TCP.
4. Heartbeat mechanism: SCTP provides the heartbeat mechanism at the network layer.

5. Security: Four-way handshake and cookie mechanisms effectively prevent the DoS attack.

As shown in Figure 23-121, the IP protocol is used at the network layer, SCTP the transport layer, and H.248/SIP the application layer.

Figure 23-121 Protocol architecture of H.248/SIP over STCP

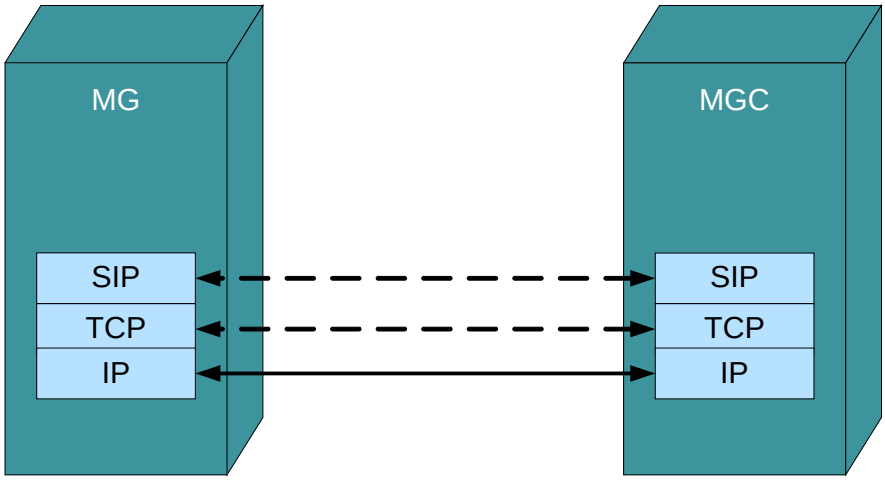


23.16.6 SIP over TCP

Some carriers require the TCP-based SIP signaling transmission, which implements the packetization of the SIP packet (the SIP packet is large in size) and enhances the transmission reliability through TCP.

As shown in Figure 23-122, the IP protocol is used at the network layer, TCP the transport layer, and SIP the application layer.

Figure 23-122 Protocol architecture of SIP over TCP



23.16.7 Voice QoS

The voice service requires high real-time performance, low delay, and fast call connection. Therefore, the voice packets should be forwarded with a high priority. The router, however, forwards the packets based on the VLAN priority (complying with 802.1p) and DSCP/ToS set in the packets.

802.1p Priority (Separately Set for Signaling and Media Streams)

Figure 23-123 802.1q frame format

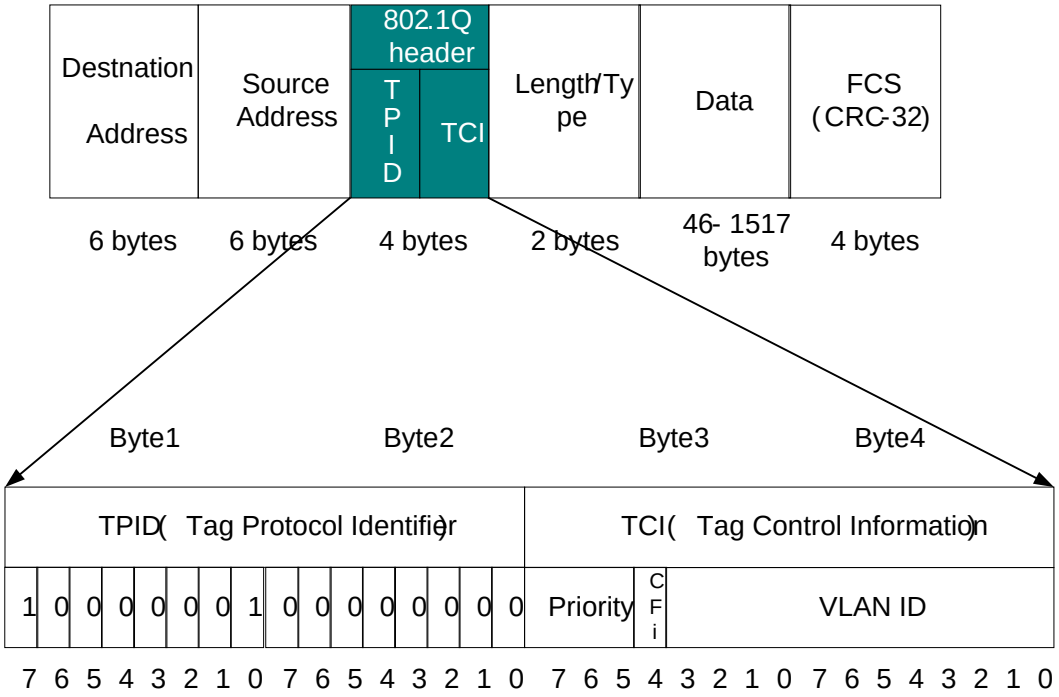


Figure 23-123 shows the Ethernet frame format defined in 802.1q. The four-byte 802.1q header contains the following contents:

- Tag protocol identifier (TPID): Two-byte tag protocol identifier, with the value of 8100.
- Tag control information (TCI): Two-byte tag control information. It is a new type of information defined by IEEE, indicating a text added with the 802.1q label. The TCI is divided into the following three fields:
  - VLAN ID: 12-bit, indicating the VLAN ID. Up to 4096 VLANs are supported. All the data packets transmitted from the host that supports 802.1q contain this field, indicating the VLAN to which the data packets belong.
  - Canonical format indicator (cfi): one-bit. It is used in the frame for data exchange between the Ethernet network of the bus type and the FDDI or token ring network.
  - Priority: three-bit, indicating the priority of the frame. Up to eight priorities are supported. It determines the data packet to be transmitted first in case of switch congestion.

The local media IP address and signaling IP address of the MA5600T/MA5603T/MA5608T can be configured in one VLAN or different VLANs according to the networking requirements. The 802.1p priorities (in the range of 0-7) can be set for the media IP address

and signaling IP address respectively. By default, the priority for either the media IP address or the signaling IP address is 6.

DSCP/TOS

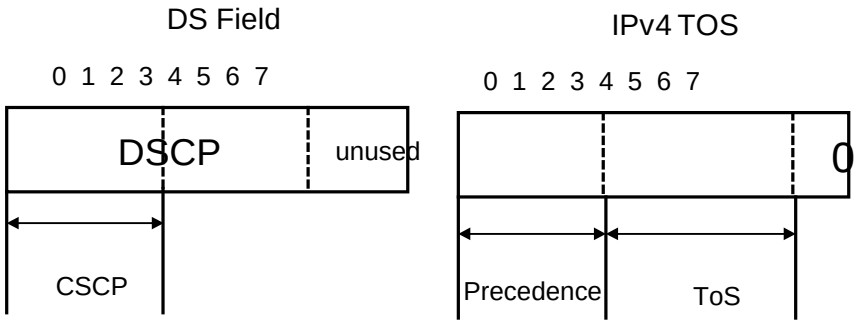
As defined in the IP protocol, the DSCP and ToS occupy the same field (one-byte) in the IP header. The device on the IP bearer network identifies whether DSCP or ToS is filled in the IP header, and schedules and forwards the packets with the DSCP/ToS field according to the settings to ensure the QoS for different services.

The type of service (ToS) field contains a three-bit precedence subfield (ignored currently), a four-bit ToS sub field, and one reserved bit (it must be set to 0). The four bits in the ToS sub field represent the minimum delay, maximum throughput, maximum reliability, and minimum cost respectively. Only one of the four bits can be set. If all four bits are set to 0, it indicates the common service.

The DSCP identification is based on the IPv4 ToS and the IPv6 traffic class.

As shown in Figure 23-124, the first six bits in the DS field (bits 0-5) are used to differentiate the DS codepoints (DSCPs) and the last two bits (bits 6 and 7) are reserved. The first three bits in the DS field (bits 0-2) are the class selector codepoint (CSCP), which indicates a class of DSCP.

Figure 23-124 DSCP identification format



DSCP is used to select the corresponding per-hop behavior (PHB) on all the nodes of the network. The PHB describes the external visible behaviors when the DS node functions on the data stream aggregation. Currently, IETF defines three types of PHB: expedited forwarding (EF), assured forwarding (AF), and best-effort. For example,

- BE: DSCP = 000000
- EF: DSCP = 101110
- The AF codepoints are as follows:

	Low Discard Priority, j = 1	Middle Discard Priority, j = 2	High Discard Priority, j = 3
AF (i = 4)	100010	100100	100110
AF (i = 3)	011010	011100	011110
AF (i = 2)	010010	010100	010110
AF (i = 1)	001010	001100	001110

The first three bits (bits 0-2) for one type of AFs are the same. To be specific, the first three bits of AF1 are 001, AF2 010, AF3 011, and AF4 100. Bits 3-4 represent the discard priority, namely, 01, 10, and 11. The larger the value, the higher the discard priority.

The DSCP/ToS value of local media IP packet and signaling IP packet can be configured on the MA5600T/MA5603T/MA5608T respectively. First the configuration policy (DSCP or ToS) is selected, and then the corresponding value is set. By default, DSCP is selected on the MA5600T/MA5603T/MA5608T, with the value of 56 (EF with the highest priority).

23.16.8 Emergency Call

The MA5600T/MA5603T/MA5608T identifies emergency calls according to emergency call digitmaps and ensures successful emergency calls by reserving digital signal processing (DSP) resources and implementing CPU overload control (OLC).

Definition

An emergency call is a kind of voice service initiated by a user using a terminal in case of emergencies. Generally, a fixed emergency call number is used within a country or region, for example, 911 in America and 110/119 in mainland China. When a user dials such a number, the system identifies the call as an emergency call, implements special processing on the call, and sends a request for routing the call to a specific emergency call center.

Device Implementation

The MA5600T/MA5603T/MA5608T identifies emergency calls according to emergency call digitmaps and ensures successful emergency calls by reserving DSP resources and implementing CPU OLC.

Digitmap

A digitmap is a set of digit collection descriptors. It is a dialing scheme resident in the MSAN and is used for detecting and reporting digit events received on a termination.

An emergency call digitmap is a dialing scheme for emergency calls. Using such digitmaps, emergency calls can be distinguished from common calls. The MSAN and core network devices, such as SoftX3000 or IMS, will then take protective measures for these emergency calls.

The MA5600T/MA5603T/MA5608T supports manual configuration of emergency call digitmaps, meeting requirements in different countries or regions.

NOTE

Digitmap configuration is complex because a digitmap is constituted by characters that have specific meanings and usage. Such information is defined in protocols and will not be detailed in this topic. It is recommended that you refer to the digitmap description in ITU-T H248.1 (applicable to the H.248 protocol) or ITU-T SIP.1 (applicable to the SIP protocol) before configuring a digitmap.

DSP Resource Reservation

DSP resources are the most important voice resources in VoIP and are used for converting TDM service flows into IP service flows. To ensure a high success rate of emergency calls, the system reserves some DSP resources (which are configurable) for callers and callees of emergency calls.

Figure 23-125 illustrates how the system reserves DSP resources for callers of emergency calls.

Figure 23-125 Reserving DSP resources for callers of emergency calls

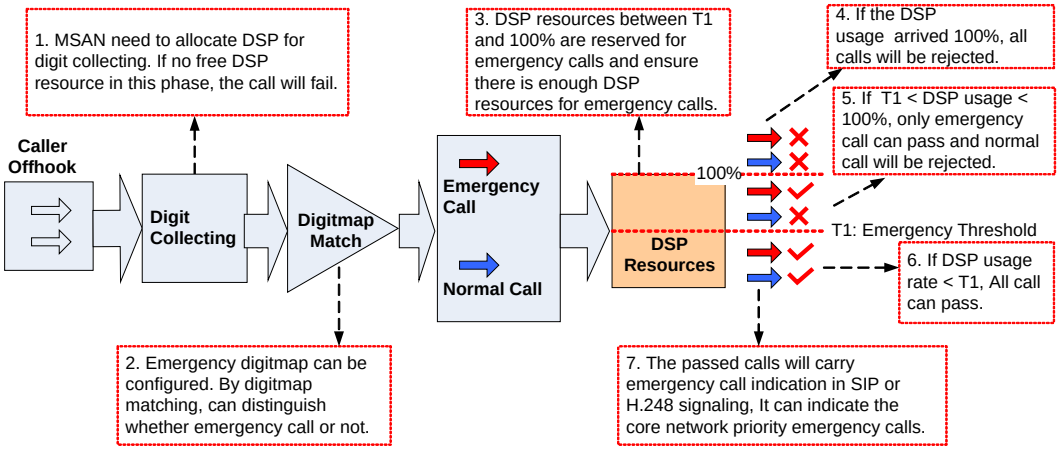
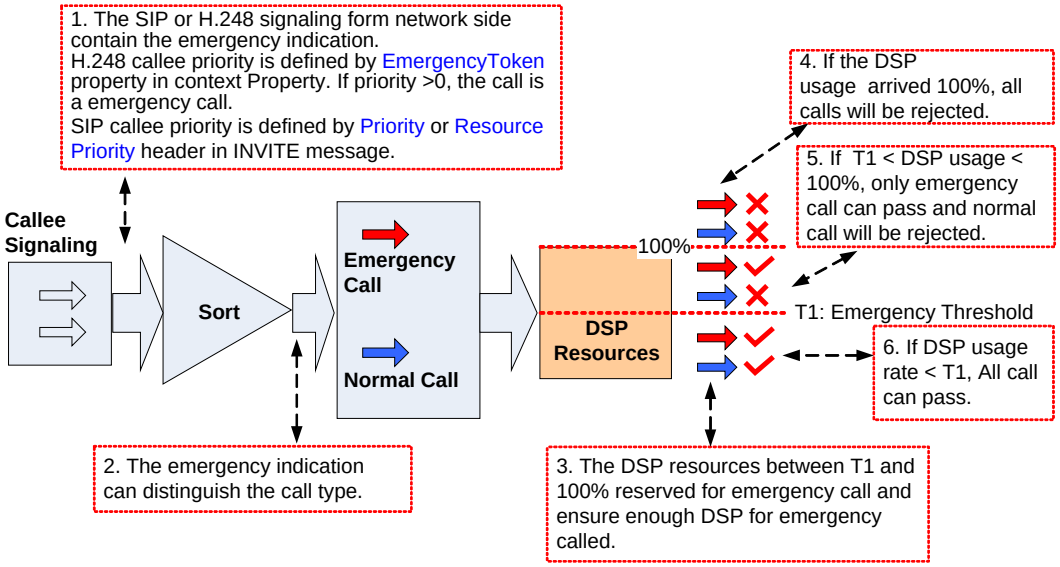


Figure 23-126 illustrates how the system reserves DSP resources for callees of emergency calls.

Figure 23-126 Reserving DSP resources for callees of emergency calls



CPU OLC

CPU OLC prevents exhaustion of equipment CPU resources. It protects equipment from service interruption or NMS management failure that is triggered by CPU overload in case of heavy traffic. CPU OLC also ensures to a certain extent the quality of high priority services when the system is overloaded.

The system employs CPU OLC for callers and callees of emergency calls, ensuring a high success rate of emergency calls.

Figure 23-127 illustrates how the system implements CPU OLC for callers of emergency calls.

Figure 23-127 Implementing CPU OLC for callers of emergency calls

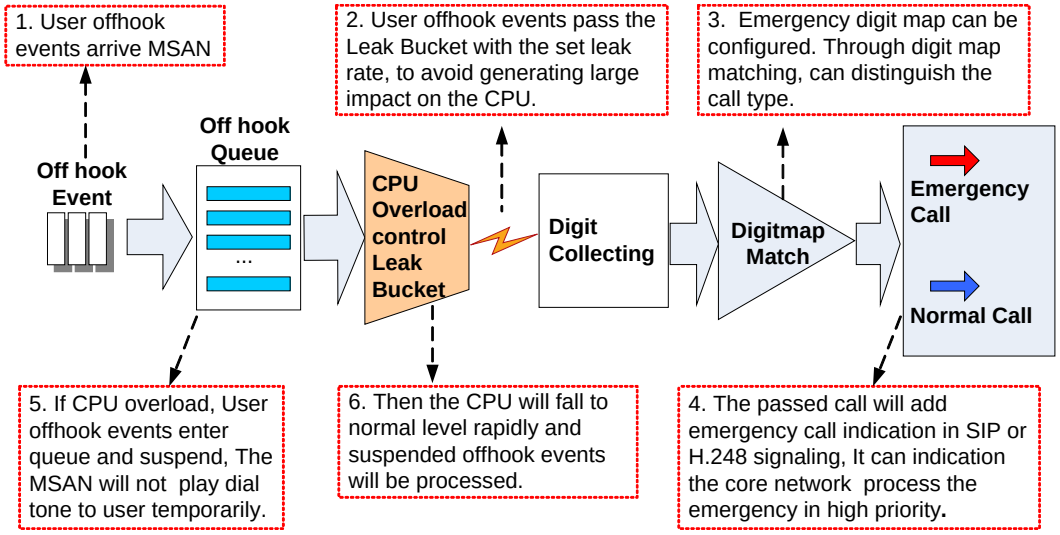
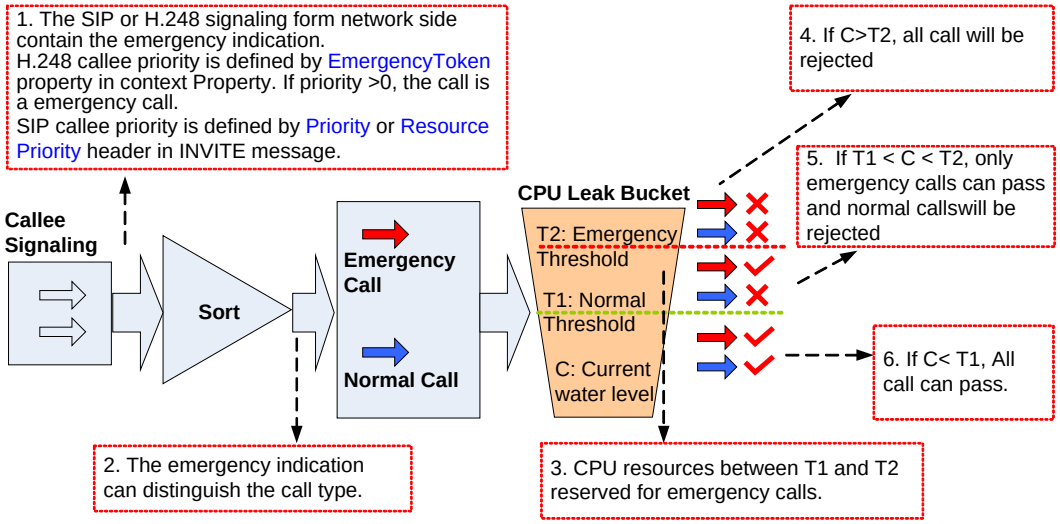


Figure 23-128 illustrates how the system implements CPU OLC for callees of emergency calls.

Figure 23-128 Implementing CPU OLC for callees of emergency calls



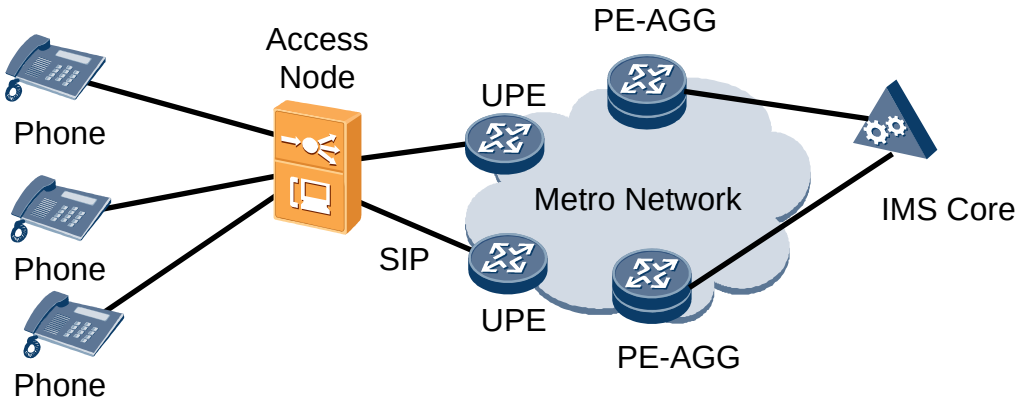
23.17 Configuring the VoIP PSTN Service (SIP-based)

The SIP-based VoIP technology makes the transport network evolve to the IP network without decreasing the voice quality, provides more value-added functions for users, and saves expense.

Application Context

As shown in Figure 23-129, the Access node functions as an 23.5.1 What Is the SIP Protocol access gateway. In the downstream direction, it provides the access to PSTN users; in the upstream direction, it is connected to the IMS system, working with the IMS core to provide the VoIP service based on SIP.

Figure 23-129 Example network of the SIP voice service



Prerequisite

- The current system protocol is the SIP protocol. If the current system protocol is not the SIP protocol, change the current system protocol to the SIP protocol with reference to the Adding an SIP Interface.
- According to the actual network, a route from the Access node to the IMS core network device must be configured to ensure that the Access node and the IMS core network device are reachable from each other.
- The voice daughter board on the control board works in the normal state.
- Electronic switch 1 must be in location-0 (indicating that the VoIP service is supported) If the SCUB control board is used. For details of the configuration method, see **electro-switch**.

Data preparation

Table 23-24 provides the data plan for configuring the VoIP service.

Table 23-24 Data plan for configuring the SIP-based VoIP service

Item			Remarks
SIP interface data (Must be the same as that on the IMS core network device.)	Parameters related to the media stream and the signaling flow	Media and signaling upstream VLAN	It is used as the upstream VLAN of the VoIP service to be configured. Note that the media stream and the signaling stream can use the same VLAN or different VLANs. The result is determined according to the negotiation with the upstream device.
		Signaling	Upstream port for configuring the SIP

Item			Remarks
		upstream port	signaling.
		Media IP address and signaling IP address	These IP addresses must be selected from the media and signaling IP address pools. The media and signaling IP address pools consists of all the IP addresses of the Layer 3 interface of the media and signaling upstream VLAN.
		Default IP address of the MG	Next hop address from the Access node to the IMS core network device. **NOTICE** If the media IP address and the signaling IP address are different and the media and the signaling are transmitted upstream through different gateways, ensure that they correspond to correct gateways. Otherwise, normal calls may not be made.
	Parameters of the SIP interface **NOTE** Parameters listed here are mandatory, which means that the SIP interface fails to be enabled if these parameters are not configured.	SIP interface ID	It is SIP interface ID used for the VoIP service to be configured, which determines the virtual access gateway (VAG) specified for the user.
		Signaling port ID of the SIP interface	The value range is 5000-5999. The protocol defines the port ID as 5060.
		IP address of the active IMS core network device to which the SIP interface belongs	When dual homing is not configured, parameters of only one IMS core network device are required. If dual homing is configured, the IP address and the port ID of the standby IMS core network device must be configured.
		Port ID of the active IMS core network device to which the SIP interface belongs	
		Transmission mode of the SIP interface	The transmission mode is selected according to the requirements on the IMS core network device. Generally, UDP is adopted.
		Home domain of the SIP interface	It corresponds to parameter **home-domain** in the MG interface attributes.
		Index of the profile used by the SIP interface	It corresponds to parameter **Profile-index** in the MG interface attributes.
		IP address obtaining mode	• In the IP mode, the IP address and the port ID of the active proxy

Item			Remarks
		of the proxy server	server must be configured. • In the DNS-A or DNS-SRV mode, the domain of the active proxy server must be configured.
	(Optional) Configuring the Ringing Mode of an MG Interface		According to the service requirements, the ringing mode of the SIP interface is determined.
Voice user data (Must be the same as that on the IMS core network device.)	Slot for the voice service board		-
	User configuration data	Phone number	The phone number that the IMS core network device allocates to the user must be configured.
		User priority	According to the service requirements, user priority needs to be specified. The user priority includes the following: • cat1: government1 (category 1 government users) • cat2: government2 (category 2 government users) • cat3: common (common users)
		User type	According to the service requirements, user type needs to be specified. The user type includes the following: • DEL: direct exchange lines (default) • ECPBX: earth calling PBX • LCPBX: loop calling PBX • PayPhone: pay phone
	(Optional) Configuring the System Parameters		The system parameters including the overseas version flag and message waiting indication (MWI) mode need to be configured according to the local standards to ensure that the response of the user terminal complies with the local standards.
	(Optional) Configuring the Overseas Parameters		The attributes such as the upper and lower thresholds of the flash-hooking duration need to be configured according to the local standards to ensure that the response of the user terminal complies with the local standards.
	(Optional) Configuring the Attributes of a PSTN Port		If the PSTN port needs to support the polarity reversal accounting, the PSTN port needs to be configured to support the polarity reversal pulse. Other

Item		Remarks
		attributes do not need to be modified if there is no special requirement.
	(Optional) Configuring the Attributes of the Ringing Current	You can adjust the ringing volume by modifying the attributes of the ringing current. Generally, the parameters of the ringing current attributes do not need to be modified. You do not need to modify the parameters of the ringing current attributes according to the local standards only when the default ringing current attributes do not meet the local standards.

Procedure

23.17.1 Configuring an SIP Interface

As an interface used for the intercommunication between the MA5600T/MA5603T/MA5608T and the MGC, the interface is vital to the SIP-based VoIP service. Therefore, to implement the VoIP service, the SIP interface must be configured and must be in the normal state.

Procedure

Configuring the Upstream VLAN Interface

This topic describes how to specify the upstream VLAN interface for the media stream and the signaling flow, and how to configure the IP addresses of the Layer 3 interface. These IP addresses are the sources of the media and signaling IP address pools.

Context

Multiple IP addresses can be configured for the same VLAN Layer 3 interface. Only one IP address functions as the primary address, and other IP addresses function as the secondary addresses.

Procedure

Run the **vlan** command to add an upstream VLAN for the media stream and the signaling flow.

Step 1 Run the **port vlan** command to add the upstream ports to the VLAN.

Step 2 Configure the IP addresses of the VLAN Layer 3 interface.

1. Run the **interface vlanif** command to enter the Layer 3 interface of the upstream VLAN for the media stream and the signaling flow.
2. Run the **ip address** command to configure the IP addresses of the Layer 3 interface.

Step 3 Run the **display interface vlanif** command to check whether the IP addresses of the Layer 3 interface are the same as those in the data plan.

----End

Example

Assume that the media stream and the signaling stream are transmitted upstream through upstream port 0/19/0. To create media and signaling upstream VLAN 3 and configure IP addresses 10.13.4.116/16 and 10.13.4.117/16 of the Layer 3 interfaces for the media IP address pool and the signaling IP address pool respectively, do as follows:

```
huawei(config)#vlan 3 standard
huawei(config)#port vlan 3 0/19 0
huawei(config)#interface vlanif 3
huawei(config-if-vlanif3)#ip address 10.13.4.116 16
huawei(config-if-vlanif3)#ip address 10.13.4.117 16 sub
huawei(config-if-vlanif3)#quit
huawei(config)#display interface vlanif 3
vlanif3 current state : UP
Line protocol current state : UP
Description : HUAWEI, SmartAX Series, vlanif3 Interface
The Maximum Transmit Unit is 1500 bytes
Internet Address is 10.13.4.116/16
Internet Address is 10.13.4.117/16 Secondary
IP Sending Frames' Format is PKTFMT_ETHNT_2, Hardware address is 00e0-fc32-1118
```

Configuring the Media and Signaling IP Address Pools

The media IP address pool defines all media IP addresses that can be used by the AG, and the signaling IP address pool defines all signaling IP addresses that can be used by the AG.

Prerequisites

The IP address of the Layer 3 interface of the media and signaling upstream VLAN must be configured. For details about how to configure the IP address, see Configuring the Upstream VLAN Interface.

Context

- The media IP address and the signaling IP address for the MG interface must be selected from the IP address pools configured here.
- The signaling IP address pool is used to store the IP addresses of the MG interfaces, and the media IP address pool is used to store the IP addresses of the media streams controlled by the signaling.
- The media IP address pool and the signaling IP address pool can be the same. Similarly, the media IP address and the signaling IP address can be the same.

NOTICE

The MGCP interface on the MA5600T/MA5603T/MA5608T does not support the isolation of media streams and signaling flows. Therefore, when the MGCP protocol is used, the media IP address pool and the signaling IP address pool must be the same. When the H.248 or SIP protocol is used, the media IP address pool and the signaling IP address pool can be the same or different.

Procedure

Run the **voip** command to enter the VoIP mode.

Step 1 Configure the media IP address pool.

1. Run the **ip address media** command to add the media IP address to the media IP address pool.

   The media IP address needs to be selected from the IP addresses of the Layer 3 interface of the media and signaling upstream VLAN.

2. Run the **display ip address media** command to check whether the media IP address pool is the same as that in the data plan.

Step 2 Configure the signaling IP address pool.

1. Run the **ip address signaling** command to add the signaling IP address to the signaling IP address pool.

   The signaling IP address needs to be selected from the IP addresses of the Layer 3 interface of the media and signaling upstream VLAN.

2. Run the **display ip address signaling** command to check whether the signaling IP address pool is the same as that in the data plan.

----End

Example

To add IP addresses 10.13.4.116/16 and 10.13.4.117/16 of Layer 3 interfaces of the media and signaling upstream VLAN to the media IP address pool and the signaling IP address pool respectively, do as follows:

```
huawei(config)#voip
huawei(config-voip)#ip address media 10.13.4.116 10.13.0.1
huawei(config-voip)#ip address media 10.13.4.117 10.13.0.1
huawei(config-voip)#ip address signaling 10.13.4.116
huawei(config-voip)#ip address signaling 10.13.4.117
huawei(config-voip)#display ip address media
  Media:
  IP Address...........: 10.13.4.116
  Subnet Mask..........: 255.255.0.0
  Gateway..............: 10.13.0.1
  MAC Address..........: 00-E0-FC-AF-91-33

  IP Address...........: 10.13.4.117
  Subnet Mask..........: 255.255.0.0
  Gateway..............: 10.13.0.1
```

```
  MAC Address..........: 00-E0-FC-AF-91-33
huawei(config-voip)#display ip address signaling
  Signaling:
  IP Address...........: 10.13.4.116
  Subnet Mask..........: 255.255.0.0
  MAC Address..........: 00-E0-FC-AF-91-33

  IP Address...........: 10.13.4.117
  Subnet Mask..........: 255.255.0.0
  MAC Address..........: 00-E0-FC-AF-91-33
```

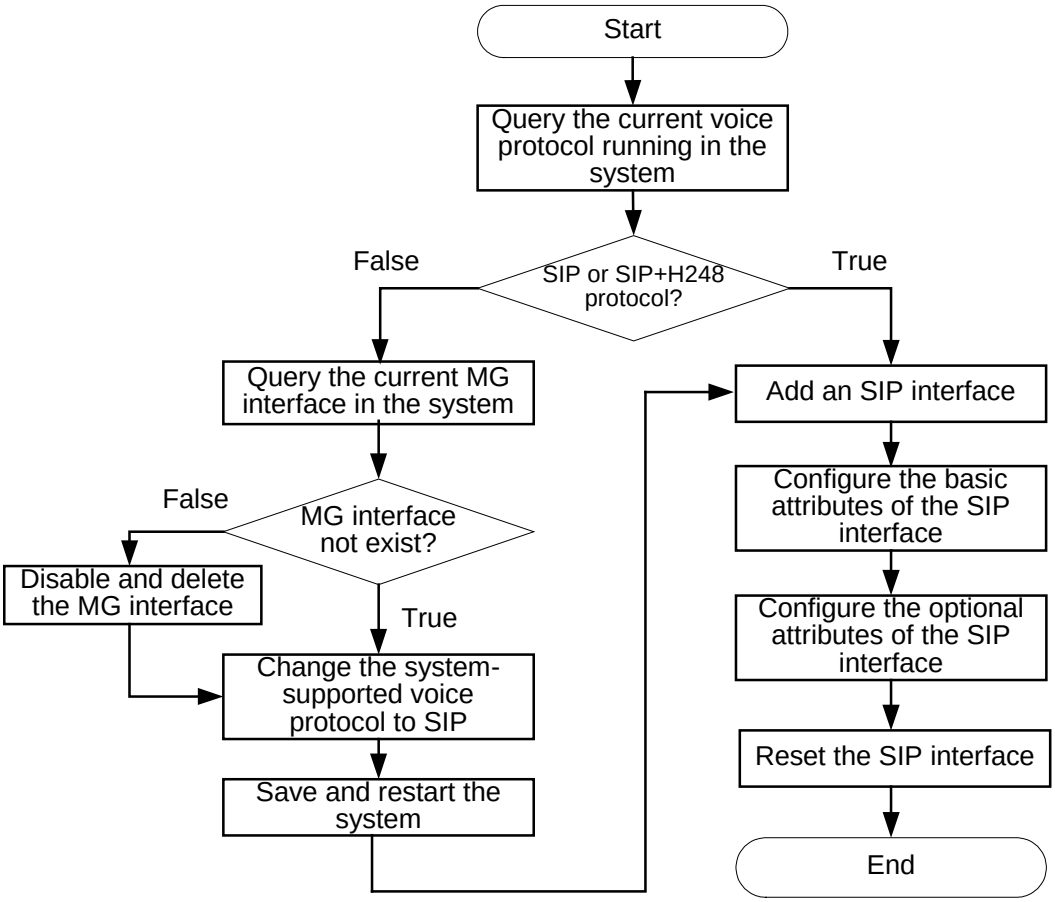
Adding an SIP Interface

The MA5600T/MA5603T/MA5608T exchanges the signaling and protocol packets with the information management system (IMS) through an SIP interface. To ensure uninterrupted signaling and protocol packet exchange, ensure that the status of the SIP interface is in a normal state after you add it.

Context

- One MA5600T/MA5603T/MA5608T supports up to eight SIP interfaces. Each SIP interface can be configured with the interface attributes separately.
- The SIP attributes configured for an SIP interface take effect on this interface only.
- The attributes of an SIP interface, including the signaling IP address, media IP address, transport-layer protocol, port ID of the transport-layer protocol, IP address (or domain name) of the proxy server, port ID of the proxy server, home domain name, profile index, are mandatory. In addition, the attributes of an SIP interface must be consistent with those configured on the IMS side so that the status of the SIP interface is normal.

Configuration Flowchart



Procedure

Query the current voice protocol running in the system.

Run the **display protocol support** command to query the voice protocol that is currently supported by the system.

- If the system voice protocol is the SIP protocol, go to Step 6.
- If the system voice protocol is not the SIP protocol, go to Step 2.

Step 1 Query the current MG interface in the system.

Run the **display if-mgcp all** or **display if-h248 all** command to query whether an MGCP interface or an H.248 interface currently exists in the system.

- If there is no such an MG interface, go to Step 4.
- If there is such an MG interface, go to Step 3.

Step 2 Disable and delete the MG interface.

1. Delete the user data of this MG interface, and then run the **shutdown(mgcp)** or **shutdown(h248)** command to disable the MG interface according to the protocol type of the interface.

NOTICE

This operation interrupts all the ongoing services carried on the MG interface. Hence, exercise caution when performing this operation. Before performing this operation, you must check whether the operation is allowed.

2. Run the **undo interface mgcp** or **undo interface h248** command to delete the MG interface.

Step 3 Change the system-supported voice protocol to SIP.

Run the **protocol support** command to change the system-supported voice protocol to SIP.

Step 4 Save the configuration data and restart the system.

Save the configuration data by running the **save** command. Then run the **reboot** command to restart the system to make the new configuration data take effect.

Step 5 Run the **interface sip** command to add an SIP interface.

Step 6 Run the **if-sip attribute basic** command to configure the basic attributes of the SIP interface.

NOTE

- Between **proxy IP** and **proxy domain**, which are basic attributes, at least one attribute must be configured. If both attributes are configured, the system determines which attribute is used according to the configured address mode of the proxy server.
- The profile index must be configured.

Step 7 Run the **if-sip attribute optional** command to configure the optional attributes of the SIP interface.

The optional attributes include some new service types supported by the SIP interface, such as the conference call, message waiting indicator, UA-profile subscription, and REG-state subscription, and also include the SIP proxy configuration mode. The optional attributes require the IMS-side support, and must be consistent with those on the IMA-side.

After the configuration is completed, run the **display if-sip attribute config** command to query the attributes of the SIP interface.

Step 8 Reset the SIP interface.

Run the **reset** command to reset the SIP interface for the new configuration data to take effect. Otherwise, the configuration data does not take effect but is only stored in the database.

After the SIP interface is reset successfully, run the **display if-sip attribute running** command to query the running status of the SIP interface. If the active (or standby) proxy is **up**, the SIP interface is normal.

----End

Example

Assume that: the media IP address is 10.10.10.13, signaling IP address is 10.10.10.13, transmission protocol is UDP, port ID is 5000, active proxy server IP address 1 is 10.10.10.14, port ID of the active proxy server is 5060, active proxy server domain name is proxy.domain, standby proxy server IP address 1 is 10.10.10.15, port ID of the standby proxy server is 5060,

home domain name is sip.huawei.com, and profile index is 1. To configure the attributes of SIP interface 0, do as follows:

```
huawei(config)#interface sip 0
huawei(config-if-sip-0)#if-sip attribute basic media-ip 10.10.10.13 signal-ip 10
.10.10.13 signal-port 5000 transfer udp primary-proxy-ip1 10.10.10.14 primary-pr
oxy-port 5060 primary-proxy-domain proxy.domain secondary-proxy-ip1 10.10.10.15
secondary-proxy-port 5060 home-domain sip.huawei.com sipprofile-index 1
huawei(config-if-sip-0)#reset
  Are you sure to reset the SIP interface?(y/n)[n]:y
huawei(config-if-sip-0)#
  Resetting SIP interface 0 succeeded
huawei(config-if-sip-0)#display if-sip attribute running
  ------------------------------------------------------------
  ...//The rest information in response to this command is omitted.
  Primary Proxy State               up   //Indicates that the SIP interface is in the
normal state.
  Secondary Proxy State              down
  ...
  ------------------------------------------------------------
```

(Optional) Configuring the Software Parameters of the SIP Interface

The software parameters of a SIP interface mainly define certain common service attributes of the SIP interface. After the software parameter configuration, the parameters take effect immediately and are valid only to the SIP interface. Skip this topic if the system default configuration meets user requirements.

Prerequisites

The MG interface has been configured. For details about how to configure the MG interface, see 23.17.1 Configuring an SIP Interface.

Context

The details about the software parameters that can be configured of a SIP interface that supports SIP, see section **Usage Guidelines** in **mg-software parameter**. The other parameters are reserved in the system.

Table 23-25 lists parameters that are usually configured to a non-default value. The other parameters are not required.

Table 23-25 Software parameters usually configured of a SIP interface

Parameter	Description	Default Setting
2	Indicates whether the standalone mode is supported.	Numeral type. Range: 0-1. • 0: indicates that the standalone function is not supported. • 1: indicates that the standalone function is

Parameter	Description	Default Setting
		supported. Default: 0 This parameter is usually set to **1**.
8	Indicates whether the heartbeat message of the MG is disabled.	Numeral type. Range: 0-1. • 0: the heartbeat message of the MG is disabled • 1: the heartbeat message of the MG is enabled Default value: 0.

Procedure

Enter the SIP interface mode.

In global config mode, run the **interface sip** command to enter the SIP interface mode.

Step 1 Configure software parameters.

Run the **mg-software parameter** command the software parameters required in the data plan.

Step 2 Check whether the software parameters are the same as those in the data plan.

Run the **display mg-software parameter** command to check whether the software parameters of the MG interface are the same as those in the data plan.

----End

Example

To configure software parameter 2 of the SIP interface 0 to 1 so that the SIP interface supports standalone, do as follows:

```
huawei(config)#interface sip 0
huawei(config-if-sip-0)#mg-software parameter 2 1
huawei(config-if-sip-0)#display mg-software parameter 2
  -------------------------------------------------
  MGID:0         para index:2   value:1
  -------------------------------------------------
 APPENDIX:
  -------------------------------------------------
 Parameter Index:  Interface software parameter name:
    2 : SAL Support
        0: No
        1: Yes
```

(Optional) Configuring the Ringing Mode of the SIP Interface

This topic describes how to configure the ringing mode of the SIP interface to support the break-make ratios of various new ringing modes and make the ringing mode meet the local standards.

Prerequisites

The SIP interface must be added successfully.

Context

- If the preset ringing modes of the system can meet the user requirements, you can select the required ringing mode and configure the corresponding ringing mapping.
- If the system-defined ringing modes cannot meet the user requirements, you can use the user-defined ringing mode and configure the corresponding ringing mapping.
- The user can configure the cadence duration for the user-defined ringing to form different ringing modes.
- The user-defined ringing modes are 0-15, which correspond to the cadence ringing modes 128-143 and initial ringing modes 144-159 defined by the user. For example, if the user-defined cadence ringing mode is 128, user-defined ringing mode 0 is selected. If the user-defined initial ringing mode is 144, user-defined ringing mode 0 is selected.
- The system supports up to 16 records of the ringing mode mapping.

Precaution

- The ringing mapping name must be unique on the same SIP interface.
- An index can be used for adding only one ringing mode on the same SIP interface.
- The 16 user-defined ringing modes can be modified but cannot be added.

Procedure

According to the Usage Guidelines of the **ringmode add** command, check whether the preset ringing mode in the system meets the requirement.

- If the requirement is met, proceed to **step 4**.
- If the requirement is not met, go to **step 2**.

Step 1 In the global config mode, run the **user defined-ring modify** command to configure the user-defined ringing mode.

NOTE
- To use the user-defined ringing mode, perform this step and you can define the ringing types numbered 0-15.
- After configuring the user-defined ringing mode, you need to reset the corresponding service board to make the configuration data take effect, so that the user of the service board can use the new user-defined ringing mode. If the service board has been in service for a period, evaluate the impact on the online users before resetting the service board, and then determine whether to perform this operation.

Step 2 Run the **display user defined-ring** command to query the user-defined ringing.

Step 3 Run the **interface sip** command to enter the SIP mode.

Step 4 Run the **ringmode add** command to add a ringing mapping.

Run this command to configure the ringing mode for the users of the same SIP interface. The key parameters are described as follows:

- cadencering: Indicates the cadence ringing mode. The range 128-143 of this parameter corresponds to user-defined ringing modes 0-15.
- initialring: Indicates the initial ringing mode. The range 144-159 of this parameter corresponds to user-defined ringing modes 0-15.

Step 5 Run the **display user defined-ring** command to query ringing mapping records.

----End

Example

To add such a ringing mode mapping record on SIP interface 0, assume that:

- Index of the ringing mode mapping record: 1
- Name of the ringing mode mapping record: alert-group
- Cadence ringing mode: 1
- Initial ringing mode: 4

do as follows:

```
huawei(config)#interface sip 0
huawei(config-if-sip-0)#ringmode add 1 alter-group cadencering 1 initialring 4
huawei(config-if-sip-0)#display ringmode 1
  ---------------------------------------------------------------
  MGID: 0
  Index: 1
  Ringmode-name: alter-group
  CadenceRing: Special Ring 1:2
  InitialRing: Normal Ring (FSK) 1:4
  ---------------------------------------------------------------
```

(Optional) Configuring User-Defined Signals

This topic describes how to configure user-defined signals on an SIP interface when the system default signals cannot meet customer requirements.

Prerequisites

The SIP interface is added successfully.

Context

Signals indicate the combination of different physical signals that are generated when the system needs to inform users of the call progress and call information. Physical signals include 3 types: media signals, line signals, and data signals.

- Media signals, including various signal tones (such as dial tone and busy tone), are generated by DSP chips.
- Line signals (such as ringing and polarity reversal signals) are generated by POTS boards.

- Data signals include calling line identification presentation (CLIP) signals, message waiting indicator (MWI) signals, and call cost signals. DSP chips process received pulse signals and then generate data signals.

When a signal is mapped to a specific scenario, the system generates this signal to inform users once this scenario occurs.

A signal may consist of a single signal unit. Figure 23-130 shows the signal configured for the scenario in which a user picks up the phone but does not dial a number for a long time. This signal consists of 2 signal units busy tone and howler tone. Table 23-26 lists signal attributes. For details, see"Parameter" of the **signal-unit set** command.

Figure 23-130 Signal configured for the scenario in which a user picks up the phone but does not dial a number for a long time

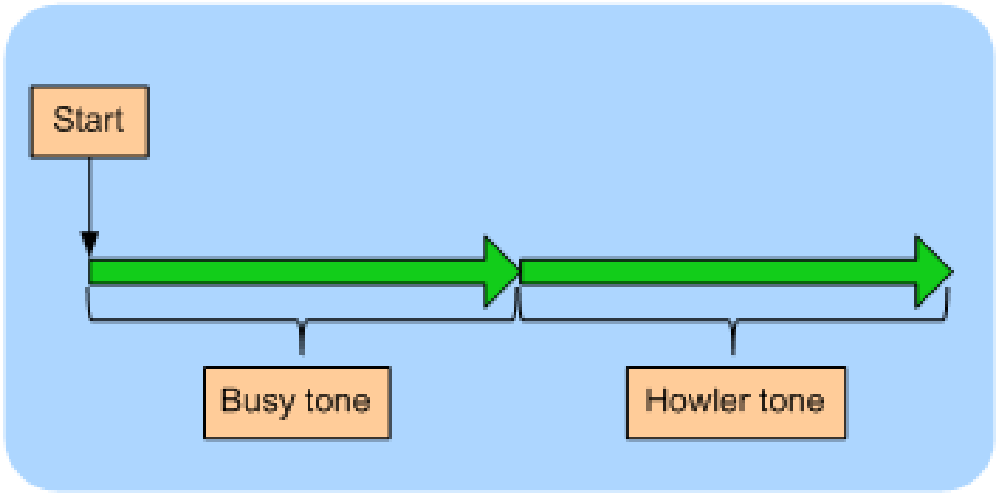


Table 23-26 Signal attributes

Attribute	Description
type	Indicates the type of the signal unit.
repeat	Indicates how many times this signal unit will be played.
duration	Indicates the duration of the signal unit. If it is set to 4294967295 (0xffffffff), the termination of signal unit playing depends on the termination condition.
start-condition	Indicates the condition for starting a signal unit.
end-condition	Indicates the condition for terminating a signal unit. If it is set to "-", this signal unit is continuously played and does not end.

Procedure

Run the **interface sip** command to enter the SIP mode.

Step 1 Run the **display signal-scene** command to query whether the system default signals meet customer requirements. If the system default signals do not meet customer requirements, configure user-defined signal by referring to **Step 3**.

Step 2 In the SIP mode, run the **signal add** command to add user-defined signals.

NOTE
The name of a user-defined signal must be unique in a system. You can run the **display signal** command to query the signals existing in the system.

Step 3 Run the **signal-unit set** command to configure signal attributes.

Step 4 Run the **signal-mapping add** command to configure the mapping between the signal and scenario.

Step 5 Run the **reset** command to reset a SIP interface for the new configuration data to take effect.

NOTE
After the SIP interface is reset successfully, you can run the **display signal-scene** command to query the new configuration data.

----End

Example

Example 1: On SIP interface 0, add a new signal named signal1 for the scenario in which a user picks up the phone but does not dial a number for a long time. This signal consists of signal unit 0 and signal unit 1.

- Type of signal unit 0: tone 1 (busy tone). Signal unit 1 will be continuously played for 40s.
- Type of signal unit 1: tone 2 (howler tone). Signal unit 2 will be continuously played for 60s.

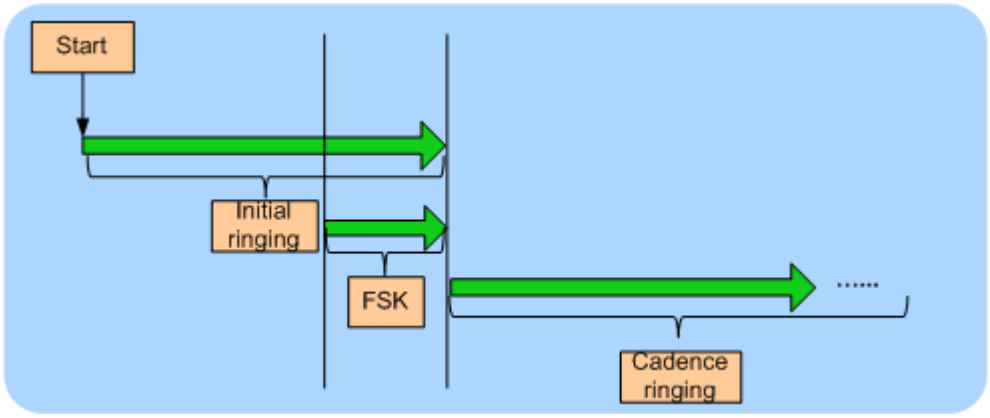
```
huawei(config)#interface sip 0
huawei(config-if-sip-0)#signal add signal1
huawei(config-if-sip-0)#signal-unit set signal1 0 type tone:1 duration 40000
huawei(config-if-sip-0)#signal-unit set signal1 1 type tone:2 duration 60000
huawei(config-if-sip-0)#signal-mapping add local:no_dial signal1
huawei(config-if-sip-0)#reset
  Are you sure to reset the SIP interface?(y/n)[n]:y
huawei(config-if-sip-0)#
  Resetting SIP interface 0 succeeded
```

Example 2: On SIP interface 0, add a new signal named signal2 for the scenario in which a user receives a normal incoming call. Figure 23-131 shows the signal configured for this scenario. This signal consists of signal unit 0, signal unit 1, and signal unit 2.

- Type of signal unit 0: init 0 (initial ringing). The initial ringing ends when signal unit 1 ends.
- Type of signal unit 1: data 0 (FSK CLIP).
  - Condition for starting signal unit 1: 650 ms elapse after the fsk_start signal is received or signal unit 0 is played for 4000 ms.

  - Condition for terminating signal unit 1: 350 ms elapse after the fsk_end signal is received. If the fsk_end signal is not received in a period of 3500 ms, signal unit 1 is automatically terminated.
- Type of signal unit 2: cade 0 (cadence ringing). When signal unit 0 ends, the system starts to play signal unit 2.

Figure 23-131 Signal2 configured for the scenario in which a user receives a normal incoming call




```
huawei(config)#interface sip 0
huawei(config-if-sip-0)#signal add signal2
huawei(config-if-sip-0)#signal-unit set signal2 0 type init:0 end-condition 1e
huawei(config-if-sip-0)#signal-unit set signal2 1 start-condition fs:650|0s:4000
duration 3500
end-condition fe:350
huawei(config-if-sip-0)#signal-unit set signal2 2 type cade:0 start-condition 0e
huawei(config-if-sip-0)#signal-mapping add local:call_incoming signal2
huawei(config-if-sip-0)#reset
  Are you sure to reset the SIP interface?(y/n)[n]:y
huawei(config-if-sip-0)#
  Resetting SIP interface 0 succeeded
```

23.17.2 Configuring the VoIP PSTN User

After an SIP interface is configured, you can add plain old telephone service (POTS) users on the SIP interface to implement the VoIP PSTN service.

Procedure

Configuring the PSTN User Data

This topic describes how to configure the PSTN user data (the same as the corresponding data on the IMS) on the SIP interface to provide the POTS terminal with the access to the network for VoIP service.

Prerequisites

The POTS service board must be inserted into the planned slot, and the RUN ALM LED of the board must be green and must be on for 1s and off for 1s repeatedly.

NOTE

You can add a service board in two ways (see the Usage Guideline of the **board add** command). It is recommended that you insert the service board into the planned slot and then confirm the board.

Procedure

In the global config mode, run the **board confirm** command to confirm the service board.

Step 1 Add a PSTN user.

1. In the global config mode, run the **esl user** command to enter the ESL user mode.
2. Run the **sippstnuser add** or **sippstnuser batadd** command to add the PSTN user.

NOTICE

- When adding a user, you can configure the phone number (parameter **telno**). When the public ID is generated by the phone number, you must enter the phone number. It is recommended that you configure this phone number the same as the phone number configured on the IMS. In addition, ensure that the phone number is unique inside the AG.

3. Run the **sippstnuser auth set** command to configure the authentication data of the PSTN user.

NOTE

Considering users safety, the IMS may require user authentication. You can run the **sippstnuser auth set** command to configure the user authentication data, including user name, password mode and password. The authentication data should be consistent with that of IMS side.

4. Run the **display sippstnuser** command to check whether the PSTN user data is the same as that in the data plan.

Step 2 (Optional) Configure the attributes of the PSTN user.

The attributes of a PSTN user need to be configured when the default configuration is not consistent with the actual application.

1. Run the **sippstnuser attribute set** or **sippstnuser attribute batset** command to configure the attributes of the PSTN user.
2. Run the **display sippstnuser attribute** command to check whether the attributes of the PSTN user are the same as those in the data plan.

----End

Example

Assume that the ASPB service board is located in slot 0/2. To configure the attributes of the 64 SIP PSTN users (phone numbers are from 83110000 to 83110063) connected to SIP interface 0, set the PSTN user type of ports from 0/2/0 to 0/2/31 to payphone, the call priorities of PSTN users from ports 0/2/32 to 0/2/63 to Cat2, and use default values for other parameters, do as follows:

```
huawei(config-esl-user)#sippstnuser batadd 0/2/0 0/2/63 0 telno 83110000 step 1
huawei(config-esl-user)#sippstnuser attribute batset 0/2/0 0/2/31 potslinetype p
ayphone
huawei(config-esl-user)#sippstnuser attribute batset 0/2/32 0/2/63 priority cat2
```

(Optional) Configuring Digitmap for SIP Interfaces

The digitmap, also called number list, refers to the dialing plan on the access gateway (AG), which is used to detect and report dialing events received at the termination point. The digitmap defines number collection rules. It allows dialing events to be reported by groups, which reduces signaling exchanges between the AG and IMS.

Prerequisites

NOTICE

The digitmap configuration is relatively complicated. The information such as the meanings and usage of the characters in a digitmap is defined in the protocol, and is not described here. This topic provides only some basic information. You are advised to refer to digitmap description in SIP standard before configuring a digitmap.

Context

- Different digitmaps are required for different services. A digitmap group includes different digitmaps, providing customized digitmaps to accommodate to users' requirements. In this way, signaling exchanges are reduced between the AG and IMS.
- A digitmap consists of digit and character strings. When the received dialing sequence matches one of the character strings, you can infer that all numbers are received.
- The priority sequence of the digitmap is: user digitmap group > interface digitmap group > global local digitmaps. If a digitmap group used by a user does not have corresponding digitmaps, this user does not have the corresponding digitmaps. For example, digitmap group A is configured in user attributes, and digitmap group B is configured in the interface of the user. Besides, two-stage out-group digitmaps are not specified in digitmap group A, but two-stage out-group digitmaps are specified in digitmap group B. When digitmaps are used, the user does not load any two-stage out-group digitmaps because digitmap group A with a highest priority does not have two-stage out-group digitmaps (although two-stage out-group digitmaps are specified in digitmap group B and local digitmaps have two-stage out-group digitmaps). If the user cannot find any user-level or interface-level digitmap groups, the user uses global local digitmaps.
- If digitmaps are not configured, the system provides a default digitmap for the user, in which all telephone numbers can be matched.

Table 23-27 provides the characters defined in the SIP protocol for digitmaps. For details, refer to the SIP standard, which provides a better guide to the digitmap configuration.

Table 23-27 SIP digitmap format

Digit or Character	Description
0-9	Indicates dialed digits 0-9.
A-D	-
E	Indicates the asterisk (*) in dual tone multiple frequency (DTMF) mode.
F	Indicates the pound key (#) in DTMF mode.
X	Indicates a wildcard, which is a digit ranging from 0 to 9.
S	Indicates the short timer. After the timer times out; that is, the dialing plan matching is complete, the system reports numbers one by one if numbers remain.
L	Indicates the long timer. After the timer times out; that is, the dialing plan matching is complete, the system reports numbers one by one if numbers remain.
Z	Indicates duration modifier, which is a dialing event with a long duration. The dialing event is located in front of the event symbol with a specified position. When the duration of the dialing event exceeds the threshold, the dialing event satisfies this position.
.	Indicates that 0 or multiple digits or characters can exist before this character.
|	Is used to isolate character strings. Each character string is a selectable dialing plan.
[]	Indicates that one of the digits or characters in the square bracket is selected.

Procedure

- Configure a digitmap.
  a. In global configuration mode, run the **local-digitmap add** command to add a local preset digitmap.
  b. (Optional) In SIP mode, run the **digitmap-timer(sip)** command to configure a digitmap timer.
- Configure a digitmap group.
  a. In global configuration mode, run the **local-digitmap add** command to add a local preset digitmap.
  b. (Optional) In SIP mode, run the **digitmap-timer(sip)** command to configure a digitmap timer.
  c. Run the **local-digitmap-group add** command to add a digitmap group.
  d. Run the **local-digitmap-group include** command to add local digitmap members to the digitmap group.

     The new digitmap group takes effect only when the user uses it in the next call.

e. Run the **mg-digitmap-group** command to configure the digitmap group used by the interface. The new digitmap group takes effect only when the user uses it in the next call.

f. Run the **sippstnuser attribute set** command to configure the digitmap group used by the user. The new digitmap group takes effect only when the user uses it in the next call.

----End

Example

For example, according to the data plan, digitmap group 1 is applied to users connected to the 0/6/0 port in the SIP interface. The digitmap group includes normal digitmaps and emergency digitmaps, whose formats are 8882xxxx and 8000xxxx respectively.

```
huawei(config)#local-digitmap add huawei normal 8882xxxx sip
huawei(config)#local-digitmap add huawei1 emergency 8000xxxx sip
huawei(config)#local-digitmap-group add DigitmapGroup1
huawei(config)#local-digitmap-group include DigitmapGroup1 huawei
huawei(config)#local-digitmap-group include DigitmapGroup1 huawei2
huawei(config)#interface sip 1
huawei(config-if-sip-1)#mg-digitmap-group DigitmapGroup1
huawei(config-if-sip-1)#quit
huawei(config)#esl user
huawei(config-esl-user)#sippstnuser attribute set 0/6/0 cliptransseq digitmap-group
DigitmapGroup1
```

(Optional) Configuring the System Parameters

This topic describes how to configure the system parameters including the overseas version flag and message waiting indication (MWI) mode according to the local standards to ensure that the response of the user terminal complies with the local standards.

Procedure

Run the **system parameters** command to configure the system parameters.

Step 1 Run the **display system parameters** command to check whether the system parameters are the same as those in the data plan.

----End

Example

To configure the overseas version flag (system parameter 1) to Hong Kong (parameter value 1), do as follows:

```
huawei(config)#system parameters 1 1
huawei(config)#display system parameters 1
  ---------------------------------------------------------------------------
  Parameter name index: 1    Parameter value: 1
  Mean: Overseas version flag, 0:China, 1:HongKong, 2:Brazil, 3:Egypt, 4:
Singapore, 5:Thailand, 6:France, 7:Britain MSFUK, 8:Britain ETSI, 9:Bulgaria,
10:Reserved, 11:Austria, 12:Hungary, 13:Poland
```

```
  ------------------------------------------------------------------------
```

(Optional) Configuring the Overseas Parameters

By default, the overseas parameters are configured according to the Chinese standards. In the actual service configuration, the attributes such as the upper and lower thresholds of the flash-hooking duration can be configured according to the local standards to ensure that the response of the user terminal complies with the local standards.

Procedure

Run the **oversea parameters** command to configure the overseas parameters.

Step 1 Run the **display oversea parameters** command to check whether the overseas parameters are the same as those in the data plan.

----End

Example

To set the upper flash-hooking threshold (overseas feature parameter 0) to 800 ms (in compliance with the Hong Kong standard) and the lower flash-hooking threshold (overseas feature parameter 1) to 100 ms (in compliance with the Hong Kong standard), do as follows:

```
huawei(config)#oversea parameters 0 800
huawei(config)#oversea parameters 1 100
huawei(config)#display oversea parameters
{ <cr>|name<U><0,99> }:

  Command:
         display oversea parameters
  ------------------------------------------------------------------------
  Parameter name index: 0    Parameter value: 800
  Mean: Hooking upper threshold(ms), reference: China:350, HongKong:800
  ------------------------------------------------------------------------
  Parameter name index: 1    Parameter value: 100
  Mean: Hooking lower threshold(ms), reference: China:100, HongKong:100
  ------------------------------------------------------------------------
  Parameter name index: 2    Parameter value: 0
  Mean: Flag of applying PARKED LINE FEED or not when user port is locked, 0:not
 apply, 1:apply
  ------------------------------------------------------------------------
  Parameter name index: 3    Parameter value: 0
  Mean: The detect time of flash upper limit to onhook, default value: 0ms
  ------------------------------------------------------------------------
  Parameter name index: 4    Parameter value: 0
  Mean: DTMF Detector Tuning, 0:Q.24, 1:Russia
------------------------------------------------------------------------
  Parameter name index: 5    Parameter value: 0
  Mean: Flag of changing TEI value for software switch, 0:no change, 1:set TEI
value to 1, 2:set TEI value to 0. Default: 0
  ------------------------------------------------------------------------
```

(Optional) Configuring the Attributes of a PSTN Port

This topic describes how to configure the attributes of a PSTN port to ensure that the PSTN port can meet the actual application requirements.

Context

The MA5600T/MA5603T/MA5608T supports the following attributes of a PSTN port:

- Physical attributes (including whether to support the polarity reversal pulse, whether to support the port locking, and dialing mode). For details about how to configure the physical attributes of a PSTN port, see **pstnport attribute set**.
- Electrical attributes (including the impedance and the current). For details about how to configure the electrical attributes of a PSTN port, see **pstnport electric set**.
- KC attributes (including the KC accounting mode and the valid voltage). For details about how to configure the KC attributes of a PSTN port, see **pstnport kc set**.

Procedure

In the global config mode, run the **pstnport** command to enter the PSTN port mode.

Step 1 Run the **pstnport attribute batset** or **pstnport attribute set** command to configure the physical attributes of the PSTN port.

Step 2 Run the **pstnport electric batset** or **pstnport electric set** command to configure the electrical attributes of the PSTN port.

Step 3 Run the **pstnport kc batset** or **pstnport kc set** command to configure the KC attributes of the PSTN port.

Step 4 Check whether the attribute configuration of the PSTN port is the same as that in the data plan.

- Run the **display pstnport attribute** command to query the physical attributes of the PSTN port.
- Run the **display pstnport electric** command to query the electrical attributes of the PSTN port.
- Run the **display pstnport kc** command to query the KC attributes of the PSTN port.

----End

Example

To configure the 32 PSTN ports of the board in slot 0/3 to support the polarity reversal accounting, do as follows:

```
huawei(config)#pstnport
huawei(config-pstnport)#pstnport attribute batset 0/3/0 0/3/31 reverse-pole-pulse
enable
huawei(config-pstnport)#display pstnport attribute 0/3
  ---------------------------------------------------------------------
  F  /S  /P            0/3 /0
  ReversePolepulse       Enable
  PulseLevel            100(ms)
  PolarityReverseMode    Hard-polarity-reverse
  Dial-Mode             DTMF-Pulse-Both
```

```
LineLock             Enable
NlpMode              Nlp normal mode
PolarityReverseWhenCLIP Disable
PulsePeriodUpperLimit   200(ms)
PulsePeriodLowerLimit   50(ms)
PulseDurationUpperLimit 90(ms)
PulseDurationLowerLimit 30(ms)
PulsePauseUpperLimit    90(ms)
PulsePauseLowerLimit    30(ms)
OffhookTime(Idle)      80(ms)
OffhookTime(Ring)      200(ms)
OffhookTime(Fsk)       50(ms)
------------------------------------------------------------------
F  /S  /P            0/3 /1
ReversePolepulse       Enable
PulseLevel            100(ms)
PolarityReverseMode    Hard-polarity-reverse
Dial-Mode             DTMF-Pulse-Both
LineLock              Enable
NlpMode               Nlp normal mode
PolarityReverseWhenCLIP Disable
PulsePeriodUpperLimit   200(ms)
PulsePeriodLowerLimit   50(ms)
PulseDurationUpperLimit 90(ms)
PulseDurationLowerLimit 30(ms)
PulsePauseUpperLimit    90(ms)
PulsePauseLowerLimit    30(ms)
OffhookTime(Idle)      80(ms)
OffhookTime(Ring)      200(ms)
OffhookTime(Fsk)       50(ms)
------------------------------------------------------------------------------
F  /S  /P            0/3 /31
ReversePolepulse       Enable
PulseLevel            100(ms)
PolarityReverseMode    Hard-polarity-reverse
Dial-Mode             DTMF-Pulse-Both
LineLock              Enable
NlpMode               Nlp normal mode
PolarityReverseWhenCLIP Disable
PulsePeriodUpperLimit   200(ms)
PulsePeriodLowerLimit   50(ms)
PulseDurationUpperLimit 90(ms)
PulseDurationLowerLimit 30(ms)
PulsePauseUpperLimit    90(ms)
PulsePauseLowerLimit    30(ms)
OffhookTime(Idle)      80(ms)
OffhookTime(Ring)      200(ms)
OffhookTime(Fsk)       50(ms)
------------------------------------------------------------------------------
```

NOTE

When a call begins and ends, the MG shows the start time and the end time based on the polarity reversal on the subscriber line. The billing terminal that supports the polarity reversal accounting function, such as a charging phone set, implements the polarity reversal accounting function based on the start time and the end time of a call.

(Optional) Configuring the Attributes of the Ringing Current

You can adjust the ringing volume and ringing tone by modifying the attributes of the ringing current. In general, the attributes of the ringing current do not need to be modified. You need to modify the attributes of the ringing current according to the local standards only when the default ringing current attributes do not meet the local standards.

Context

The attributes of the ringing current include the following two parameters:

- Ringing current frequency: A higher frequency indicates a sharper ringing tone. The default ringing current frequency is 25 Hz.
- AC amplitude (AC voltage): A greater amplitude indicates a louder ringing tone. The default AC amplitude is 75 Vrms.

Procedure

In the global config mode, run the **voip** command to enter the VoIP mode.

Step 1 Run the **ring attribute set** command to configure the attributes of the ringing current according to the data plan.

Step 2 Run the **display ring attribute** command to check whether the attributes of the ringing current are the same as those in the data plan.

----End

Example

To set the ringing current frequency to 50 Hz (parameter value 2), AC amplitude to 50 Vrms (parameter value 2), do as follows:

```
huawei(config-voip)#ring attribute set frequency 2 acamplitude 2
huawei(config-voip)#display ring attribute
  ringing current frequency  : 50HZ
  ringing current acamplitute: 50VRMS
```

23.17.3 (Optional) Configuring Line Hunting

When multiple E1 lines exist in the upstream direction of a private branch exchange (PBX), hunting is performed based on the pilot number or other accounts required by the PBX. The line hunting function allows one or multiple accounts to share a group of ports by configuring hunting groups and hunting policies.

Context

- A hunting group consists of ports, subhunting groups, and hunting rules. A sub hunting group also consists of ports, sub hunting groups, and hunting rules.
- A wildcard number can be configured for hunting groups, for example, 024545*. You can also configure a direct dialing number.
- A port can belong to multiple hunting groups which must be in the same VAG.
- Only the Session Initiation Protocol (SIP) supports the line hunting function.

Procedure

Run the **hunting-group add** command to add a hunting group. Then the hunting group is added to the specified SIP interface.

- **hunting-mode** indicates the hunting policy. This parameter can be set to **order**, **round-robin**, or **weighted-round-robin**.
  - **order**: indicates the sequential hunting.
  - **round-robin**: indicates the circular hunting.
  - **weighted-round-robin**: indicates the circular hunting by weight.

Step 1 Configure users on the SIP interface. Run different commands when configuring different users. Specifically,

- Run the **sippstnuser add** command when configuring PSTN users.
- Run the **sipbrauser add** command when configuring BRA users.
- Run the **sipprauser add** command when configuring PRA users.

Step 2 Run the **hunting-group member add** command to add hunting group members. A hunting group member can be a single port or a sub hunting group. After hunting group members are added, hunting is performed based on configurations.

Step 3 Run the **group-number add** command to add a hunting group account. After the group account is used by the hunting group, the account is called in based on hunting policies.

----End

Example

For example, when number 2878000 is dialed, hunting is cyclically performed between 0/2/1, 0/3/1, and 0/4/1 ports. When number 2878001 is dialed, the 0/2/1 is selected with preference. When the 0/2/1 port is busy, the 0/3/1 port is selected and then the 0/4/1 port. Configurations are as follows.

Parameter	HG1	HG2
hunting-mode	round-robin	order
inherit-flag	disable	disable
rotary	disable	disable
Group members	0/2/1, 0/3/1, 0/4/1	0/2/1, 0/3/1, 0/4/1

```
huawei(config)#interface sip 0
Are you sure to add the SIP interface?(y/n)[n]:y
huawei(config-if-sip-0)#hunting-group add HG1 hunting-mode round-robin inherit-flag
disable
 rotary disable
huawei(config-if-sip-0)#hunting-group add HG2 hunting-mode order inherit-flag disable
rot
ary disable
huawei(config-if-sip-0)#quit
huawei(config)#esl user
```

```
huawei(config-esl-user)#sippstnuser add 0/2/1 0
huawei(config-esl-user)#sippstnuser add 0/3/1 0
huawei(config-esl-user)#sippstnuser add 0/4/1 0
huawei(config-esl-user)#quit
huawei(config)#interface sip 0
huawei(config-if-sip-0)#hunting-group member add HG1 0/2/1 6
huawei(config-if-sip-0)#hunting-group member add HG1 0/3/1 7
huawei(config-if-sip-0)#hunting-group member add HG1 0/4/1 8
huawei(config-if-sip-0)#hunting-group member add HG2 0/2/1 6
huawei(config-if-sip-0)#hunting-group member add HG2 0/3/1 7
huawei(config-if-sip-0)#hunting-group member add HG2 0/4/1 8
huawei(config-if-sip-0)#group-number add 2878000 hunting-group HG1
huawei(config-if-sip-0)#group-number add 2878001 hunting-group HG2
```

23.18 Configuring the VoIP ISDN BRA Service (SIP-based)

After configuring the SIP interface, you can add VoIP ISDN BRA users on this interface to implement the VoIP ISDN BRA service.

Prerequisites

According to the actual network, a route from the MA5600T/MA5603T/MA5608T to the IMS must be configured to ensure that the MA5600T/MA5603T/MA5608T communicates with the IMS normally.

Context

- The ISDN is integrated services digital network. Defined by the International Telegraph and Telephone Consultative Committee (CCITT), the ISDN technology provides two types of user-network interfaces based on the user-network interface reference model.
  - ISDN basic rate interface (BRI), which supports a rate of 144 kbit/s and provides two B channels (for carrying services) and one D channel (for transmitting call control signaling and maintenance and management signaling). The rate of B channels is 64 kbit/s and the rate of D channel is 16 kbit/s. The service carried on the ISDN BRI is the ISDN basic rate access (BRA) service.
  - ISDN primary rate interface (PRI), which supports a rate of 2.048 Mbit/s and provides 30 B channels and one D channel. The rates of the B channel and D channel are both 64 kbit/s. The service carried on the ISDN PRI is the ISDN primary rate access (PRA) service.

The MA5600T/MA5603T/MA5608T can also function as a voice over IP gateway (VGW) in the IMS architecture. In the downstream direction, it is connected to the ISDN BRA users; in the upstream direction, it is connected to the IMS system through the SIP interface, providing the VoIP ISDN BRA service by working with the IMS core.

The functions and applications of the SIP interface are the same as the functions and applications of the MG interface.

Data preparation

Table 23-28 provides the data plan for configuring the VoIP ISDN BRA service.

Table 23-28 Data plan for configuring the VoIP ISDN BRA service when the SIP protocol is used

Item			Remarks
SIP interface data	Media and signaling parameters	Media and signaling upstream VLAN	It is used for the upstream VLAN of the VoIP service to be configured. **NOTICE** Note that the media and the signaling can use the same VLAN or different VLANs, depending on the negotiation with the upstream device.
		Signaling upstream port	Uplink port for configuring the SIP signaling.
		Media IP address and signaling IP addresses	These IP addresses must be selected from the media and signaling IP address pools. The media and signaling IP address pools consists of all the IP addresses of the L3 interface of the media and signaling upstream VLAN.
		Default IP address of the MG	Next hop address from the MA5600T/MA5603T/MA5608T to the IMS core network device. **NOTICE** If the media IP address and the signaling IP address are different and the media and the signaling are transmitted upstream through different gateways, ensure that they correspond to correct gateways. Otherwise, the call service may fail.
	Attributes of the SIP interface **NOTE** Parameters listed here are mandatory, which means that the SIP interface cannot be enabled if these parameters are not configured.	SIP interface ID	It is used for the VoIP service to be configured.
		Signaling port ID of the SIP interface	The value range is 5000-5999. The protocol defines the port ID as 5060.
		IP address of the active IMS core network device to which the SIP interface belongs	When dual homing is not configured, parameters of only one IMS core network device are required. When dual homing is required, the IP address and the port ID of the standby IMS core network device must be configured.
		Port ID of the active IMS core network device to which the SIP interface belongs	
		Transmission	The transmission mode is selected

Item			Remarks
		mode of the SIP interface	according to the requirements of the IMS core network device. Generally, UDP is used.
		Home domain of the SIP interface	It corresponds to parameter **home-domain** in the MG interface attributes.
		Index of the profile used by the SIP interface	It corresponds to parameter **Profile-index** in the MG interface attributes.
		IP address obtaining mode of the proxy server	• In the IP mode, the IP address and the port ID of the active proxy server must be configured. • In the DNS-A or DNS-SRV mode, the domain name of the active proxy server must be configured.
ISDN BRA user data (The data configuration must be consistent with the data configuration on the IMS.)	Slot that houses the BRA service board.		-
	Configuring the ISDN BRA User Data	Phone number	The phone number that the IMS core network device allocates to the user must be configured.
		User priority	According to the service requirements, user priorities must be specified, including: • cat1: government1 (category 1 government users) • cat2: government2 (category 2 government users) • cat3: common (default priority, namely, common users)
	(Optional) Configuring the System Parameters		The system parameters including the oversea version flag and message waiting indication (MWI) mode need to be configured according to local standards to ensure that the response of the user terminal meets local standards.
	(Optional) Configuring the Overseas Parameters		The attributes such as the upper and lower thresholds of the flash-hooking duration need to be configured according to local standards to ensure that the response of the user terminal meets local standards.
	(Optional) Configuring the Attributes of the ISDN BRA Port		The attributes such as the working mode, remote power supply status, and auto-deactivation status of the port

Item		Remarks
		can be configured. Modify such attributes only if there is a special requirement.

23.18.1 Configuring the SIP Interface

As an interface used for the intercommunication between the MA5600T/MA5603T/MA5608T and the MGC, the interface is vital to the SIP-based VoIP service. Therefore, to implement the VoIP service, the SIP interface must be configured and must be in the normal state.

Procedure

Configuring the Upstream VLAN Interface

This topic describes how to specify the upstream VLAN interface for the media stream and the signaling flow, and how to configure the IP addresses of the Layer 3 interface. These IP addresses are the sources of the media and signaling IP address pools.

Context

Multiple IP addresses can be configured for the same VLAN Layer 3 interface. Only one IP address functions as the primary address, and other IP addresses function as the secondary addresses.

Procedure

Run the **vlan** command to add an upstream VLAN for the media stream and the signaling flow.

Step 1 Run the **port vlan** command to add the upstream ports to the VLAN.

Step 2 Configure the IP addresses of the VLAN Layer 3 interface.

1. Run the **interface vlanif** command to enter the Layer 3 interface of the upstream VLAN for the media stream and the signaling flow.
2. Run the **ip address** command to configure the IP addresses of the Layer 3 interface.

Step 3 Run the **display interface vlanif** command to check whether the IP addresses of the Layer 3 interface are the same as those in the data plan.

----End

Example

Assume that the media stream and the signaling stream are transmitted upstream through upstream port 0/19/0. To create media and signaling upstream VLAN 3 and configure IP addresses 10.13.4.116/16 and 10.13.4.117/16 of the Layer 3 interfaces for the media IP address pool and the signaling IP address pool respectively, do as follows:

```
huawei(config)#vlan 3 standard
huawei(config)#port vlan 3 0/19 0
```

```
huawei(config)#interface vlanif 3
huawei(config-if-vlanif3)#ip address 10.13.4.116 16
huawei(config-if-vlanif3)#ip address 10.13.4.117 16 sub
huawei(config-if-vlanif3)#quit
huawei(config)#display interface vlanif 3
vlanif3 current state : UP
Line protocol current state : UP
Description : HUAWEI, SmartAX Series, vlanif3 Interface
The Maximum Transmit Unit is 1500 bytes
Internet Address is 10.13.4.116/16
Internet Address is 10.13.4.117/16 Secondary
IP Sending Frames' Format is PKTFMT_ETHNT_2, Hardware address is 00e0-fc32-1118
```

Configuring the Media and Signaling IP Address Pools

The media IP address pool defines all media IP addresses that can be used by the AG, and the signaling IP address pool defines all signaling IP addresses that can be used by the AG.

Prerequisites

The IP address of the Layer 3 interface of the media and signaling upstream VLAN must be configured. For details about how to configure the IP address, see Configuring the Upstream VLAN Interface.

Context

- The media IP address and the signaling IP address for the MG interface must be selected from the IP address pools configured here.
- The signaling IP address pool is used to store the IP addresses of the MG interfaces, and the media IP address pool is used to store the IP addresses of the media streams controlled by the signaling.
- The media IP address pool and the signaling IP address pool can be the same. Similarly, the media IP address and the signaling IP address can be the same.

NOTICE

The MGCP interface on the MA5600T/MA5603T/MA5608T does not support the isolation of media streams and signaling flows. Therefore, when the MGCP protocol is used, the media IP address pool and the signaling IP address pool must be the same. When the H.248 or SIP protocol is used, the media IP address pool and the signaling IP address pool can be the same or different.

Procedure

Run the **voip** command to enter the VoIP mode.

Step 1 Configure the media IP address pool.

1. Run the **ip address media** command to add the media IP address to the media IP address pool.

The media IP address needs to be selected from the IP addresses of the Layer 3 interface of the media and signaling upstream VLAN.

2. Run the **display ip address media** command to check whether the media IP address pool is the same as that in the data plan.

Step 2 Configure the signaling IP address pool.

1. Run the **ip address signaling** command to add the signaling IP address to the signaling IP address pool.

   The signaling IP address needs to be selected from the IP addresses of the Layer 3 interface of the media and signaling upstream VLAN.

2. Run the **display ip address signaling** command to check whether the signaling IP address pool is the same as that in the data plan.

----End

Example

To add IP addresses 10.13.4.116/16 and 10.13.4.117/16 of Layer 3 interfaces of the media and signaling upstream VLAN to the media IP address pool and the signaling IP address pool respectively, do as follows:

```
huawei(config)#voip
huawei(config-voip)#ip address media 10.13.4.116 10.13.0.1
huawei(config-voip)#ip address media 10.13.4.117 10.13.0.1
huawei(config-voip)#ip address signaling 10.13.4.116
huawei(config-voip)#ip address signaling 10.13.4.117
huawei(config-voip)#display ip address media
  Media:
  IP Address...........: 10.13.4.116
  Subnet Mask..........: 255.255.0.0
  Gateway..............: 10.13.0.1
  MAC Address..........: 00-E0-FC-AF-91-33

  IP Address...........: 10.13.4.117
  Subnet Mask..........: 255.255.0.0
  Gateway..............: 10.13.0.1
  MAC Address..........: 00-E0-FC-AF-91-33
huawei(config-voip)#display ip address signaling
  Signaling:
  IP Address...........: 10.13.4.116
  Subnet Mask..........: 255.255.0.0
  MAC Address..........: 00-E0-FC-AF-91-33

  IP Address...........: 10.13.4.117
  Subnet Mask..........: 255.255.0.0
  MAC Address..........: 00-E0-FC-AF-91-33
```

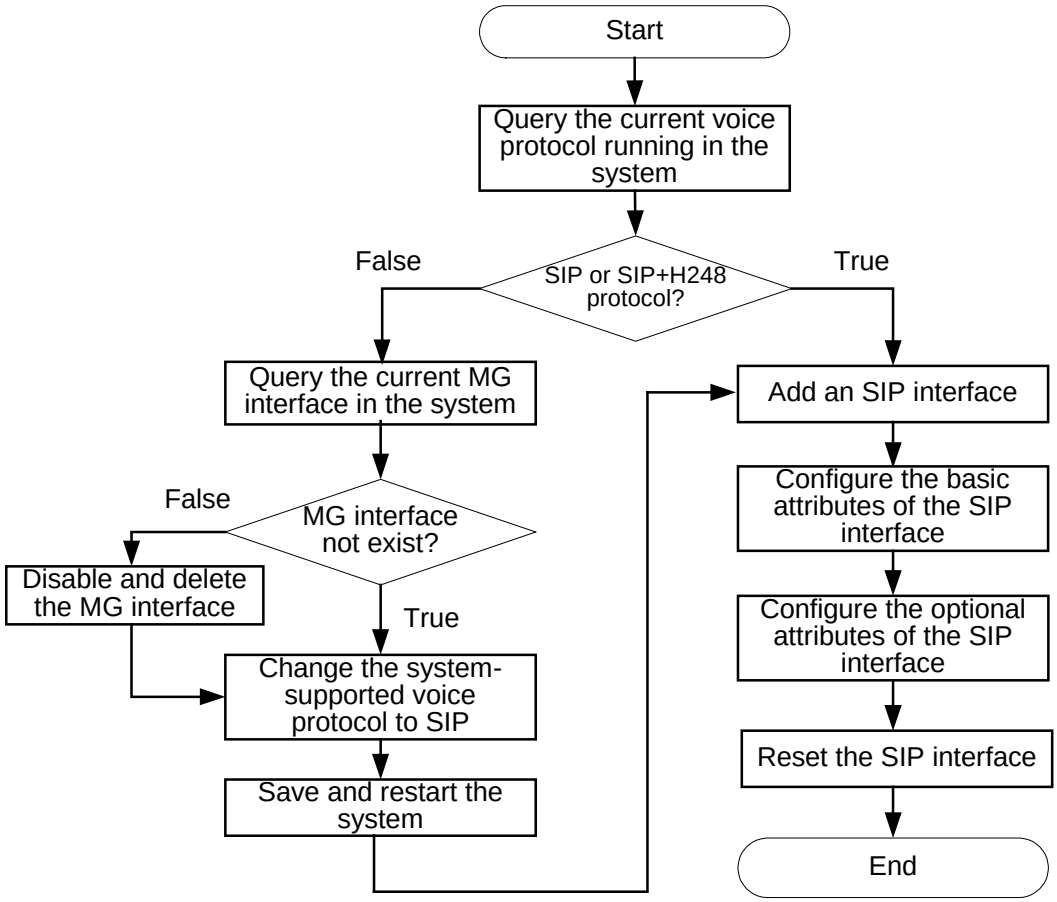
Adding an SIP Interface

The MA5600T/MA5603T/MA5608T exchanges the signaling and protocol packets with the information management system (IMS) through an SIP interface. To ensure uninterrupted signaling and protocol packet exchange, ensure that the status of the SIP interface is in a normal state after you add it.

Context

- One MA5600T/MA5603T/MA5608T supports up to eight SIP interfaces. Each SIP interface can be configured with the interface attributes separately.
- The SIP attributes configured for an SIP interface take effect on this interface only.
- The attributes of an SIP interface, including the signaling IP address, media IP address, transport-layer protocol, port ID of the transport-layer protocol, IP address (or domain name) of the proxy server, port ID of the proxy server, home domain name, profile index, are mandatory. In addition, the attributes of an SIP interface must be consistent with those configured on the IMS side so that the status of the SIP interface is normal.

Configuration Flowchart



Procedure

Query the current voice protocol running in the system.

Run the **display protocol support** command to query the voice protocol that is currently supported by the system.

- If the system voice protocol is the SIP protocol, go to Step 6.
- If the system voice protocol is not the SIP protocol, go to Step 2.

Step 1 Query the current MG interface in the system.

Run the **display if-mgcp all** or **display if-h248 all** command to query whether an MGCP interface or an H.248 interface currently exists in the system.

- If there is no such an MG interface, go to Step 4.
- If there is such an MG interface, go to Step 3.

Step 2 Disable and delete the MG interface.

1. Delete the user data of this MG interface, and then run the **shutdown(mgcp)** or **shutdown(h248)** command to disable the MG interface according to the protocol type of the interface.

NOTICE

This operation interrupts all the ongoing services carried on the MG interface. Hence, exercise caution when performing this operation. Before performing this operation, you must check whether the operation is allowed.

2. Run the **undo interface mgcp** or **undo interface h248** command to delete the MG interface.

Step 3 Change the system-supported voice protocol to SIP.

Run the **protocol support** command to change the system-supported voice protocol to SIP.

Step 4 Save the configuration data and restart the system.

Save the configuration data by running the **save** command. Then run the **reboot** command to restart the system to make the new configuration data take effect.

Step 5 Run the **interface sip** command to add an SIP interface.

Step 6 Run the **if-sip attribute basic** command to configure the basic attributes of the SIP interface.

NOTE

- Between **proxy IP** and **proxy domain**, which are basic attributes, at least one attribute must be configured. If both attributes are configured, the system determines which attribute is used according to the configured address mode of the proxy server.
- The profile index must be configured.

Step 7 Run the **if-sip attribute optional** command to configure the optional attributes of the SIP interface.

The optional attributes include some new service types supported by the SIP interface, such as the conference call, message waiting indicator, UA-profile subscription, and REG-state subscription, and also include the SIP proxy configuration mode. The optional attributes require the IMS-side support, and must be consistent with those on the IMA-side.

After the configuration is completed, run the **display if-sip attribute config** command to query the attributes of the SIP interface.

Step 8 Reset the SIP interface.

Run the **reset** command to reset the SIP interface for the new configuration data to take effect. Otherwise, the configuration data does not take effect but is only stored in the database.

After the SIP interface is reset successfully, run the **display if-sip attribute running** command to query the running status of the SIP interface. If the active (or standby) proxy is **up**, the SIP interface is normal.

----End

Example

Assume that: the media IP address is 10.10.10.13, signaling IP address is 10.10.10.13, transmission protocol is UDP, port ID is 5000, active proxy server IP address 1 is 10.10.10.14, port ID of the active proxy server is 5060, active proxy server domain name is proxy.domain, standby proxy server IP address 1 is 10.10.10.15, port ID of the standby proxy server is 5060, home domain name is sip.huawei.com, and profile index is 1. To configure the attributes of SIP interface 0, do as follows:

```
huawei(config)#interface sip 0
huawei(config-if-sip-0)#if-sip attribute basic media-ip 10.10.10.13 signal-ip 10
.10.10.13 signal-port 5000 transfer udp primary-proxy-ip1 10.10.10.14 primary-pr
oxy-port 5060 primary-proxy-domain proxy.domain secondary-proxy-ip1 10.10.10.15
secondary-proxy-port 5060 home-domain sip.huawei.com sipprofile-index 1
huawei(config-if-sip-0)#reset
  Are you sure to reset the SIP interface?(y/n)[n]:y
huawei(config-if-sip-0)#
  Resetting SIP interface 0 succeeded
huawei(config-if-sip-0)#display if-sip attribute running
  ------------------------------------------------------------
  ...//The rest information in response to this command is omitted.
  Primary Proxy State               up   //Indicates that the SIP interface is in the
normal state.
  Secondary Proxy State              down
  ...
  ------------------------------------------------------------
```

(Optional)Configuring the Software Parameters of an SIP Interface

The software parameters of a SIP interface mainly define certain common service attributes of the SIP interface. After the software parameter configuration, the parameters take effect immediately and are valid only to the SIP interface. Skip this topic if the system default configuration meets user requirements.

Prerequisites

The MG interface has been configured. For details about how to configure the MG interface, see 23.18.1 Configuring the SIP Interface.

Context

The details about the software parameters that can be configured of a SIP interface that supports SIP, see section **Usage Guidelines** in **mg-software parameter**. The other parameters are reserved in the system.

Table 23-29 lists parameters that are usually configured to a non-default value. The other parameters are not required.

Table 23-29 Software parameters usually configured of a SIP interface

Parameter	Description	Default Setting
2	Indicates whether the standalone mode is supported.	Numeral type. Range: 0-1. • 0: indicates that the standalone function is not supported. • 1: indicates that the standalone function is supported. Default: 0 This parameter is usually set to **1**.
8	Indicates whether the heartbeat message of the MG is disabled.	Numeral type. Range: 0-1. • 0: the heartbeat message of the MG is disabled • 1: the heartbeat message of the MG is enabled Default value: 0.

Procedure

Enter the SIP interface mode.

In global config mode, run the **interface sip** command to enter the SIP interface mode.

Step 1 Configure software parameters.

Run the **mg-software parameter** command the software parameters required in the data plan.

Step 2 Check whether the software parameters are the same as those in the data plan.

Run the **display mg-software parameter** command to check whether the software parameters of the MG interface are the same as those in the data plan.

----End

Example

To configure software parameter 2 of the SIP interface 0 to 1 so that the SIP interface supports standalone, do as follows:

```
huawei(config)#interface sip 0
huawei(config-if-sip-0)#mg-software parameter 2 1
huawei(config-if-sip-0)#display mg-software parameter 2
  ------------------------------------------------
  MGID:0         para index:2   value:1
  ------------------------------------------------
 APPENDIX:
  ------------------------------------------------
 Parameter Index:  Interface software parameter name:
```

```
2 : SAL Support
   0: No
   1: Yes
```

23.18.2 Configuring the VoIP ISDN BRA User

This topic describes how to configure the VoIP ISDN BRA user. After the SIP interface is configured, you can add the VoIP ISDN BRA user on this interface to implement the VoIP ISDN BRA service.

Configuring the ISDN BRA User Data

This topic describes how to configure the ISDN BRA user data on the SIP interface (the data must be the same as the corresponding data on the IMS) so that the ISDN BRA user can access the network to use the ISDN BRA service.

Prerequisites

The ISDN BRA service board must be inserted into the planned slot correctly and the board must be in the normal state.

Default Configuration

Table 23-30 lists the default settings of the attributes of the ISDN BRA user. When configuring the attributes of these attributes, you can modify the values according to the service requirements.

Table 23-30 Default settings of the attributes of the ISDN BRA user

Parameter	Default Settings
Priority of the ISDN BRA user	cat3 (common users)
Flag of reporting the UNI fault of the ISDN BRA user	disable
Threshold for the number of auto recoveries from deterioration faults	20
The matching scheme of an outgoing calling number	match
The voice bear capability of a port	speech+3.1k-audio
The type of the called number sent to the ISDN terminal when an incoming call is initiated	unknown
The international prefix flag	disable
The national prefix flag	disable
The change plan of an incoming called number used by the user on the port	-
The change plan of an outgoing calling	-

Parameter	Default Settings
number used by the user on the port	
The calling number	-
The digitmap group used by the user	-

Procedure

In global config mode, run the **esl user** command to enter the ESL user mode.

Step 1 Run the **sipbrauser add** command to add an ISDN BRA user.

Step 2 Run the **display sipbrauser** command to check whether the ISDN BRA user data is the same as the data plan.

Step 3 (Perform this step when you need to modify the attributes of an ISDN BRA user.) Run the **sipbrauser attribute set** command to configure the attributes of the ISDN BRA user.

Step 4 (Perform this step only after you modify the attributes of the ISDN BRA user.) Run the **display sipbrauser attribute** command to query whether the configured attributes of the ISDN BRA user are the same as the data plan.

Step 5 (Perform this step only when you need to configure an extended phone number for the ISDN BRA user.) Run the **sipbrauser extend-telno add** command to add multiple phone numbers or a phone number containing non-digit characters for the ISDN BRA user.

Step 6 (Perform this step only after you configure the extended phone number for an ISDN BRA user.) Run the **display sipbrauser extend-telno** command to query whether the configured extended phone number of the ISDN BRA user is the same as the data plan.

----End

Example

Assume that:

- SIP interface ID: 0
- Phone number: 28780000
- Call priority: cat3
- UNI fault report flag: disable
- Number of auto recoveries from deterioration faults: 10
- Voice bear capability: speech
- Type of the called number: national
- International prefix flag: enable
- national prefix flag: enable
- Digitmap group: DigitmapGroup1
- Change plan of an incoming called number: NumberChange1
- Change plan of an outgoing calling number: NumberChange2
- Calling number is 12345678

- Extended phone number: +86-755-28780000

To add such an ISDN BRA user connected to the 0/4/0 port, do as follows:

```
huawei(config)#esl-user
huawei(config-esl-user)#sipbrauser add 0/4/0 0 telno 28780000
huawei(config-esl-user)#display sipbrauser 0/4/0
{ <cr>|endframeid/slotid/portid<S><Length 1-15> }:

  Command:
         display sipbrauser 0/4/0
  ---------------------------------------
  F  /S /P    MGID    TelNo
  ---------------------------------------
  0/4/0    0       28780000
  ---------------------------------------
huawei(config-esl-user)#sipbrauser attribute set 0/4/0 priority cat3 unireport d
isable auto-resume-limit 10 bc speech called-num-type national international-pre
fix-flag enable national-prefix-flag enable digitmap-group DigitmapGroup1 incomi
ngcall-numberchange NumberChange1 outgoingcall-numberchange NumberChange2 cli-nu
mber 12345678
huawei(config-esl-user)#display sipbrauser attribute 0/4/0
{ <cr>|endframeid/slotid/portid<S><Length 1-15> }:

  Command:
         display sipbrauser attribute 0/4/0
  ------------------------------------------------------------------------------
  F  /S /P                      : 0/4/0
  UNIreport                     : disable
  Priority                      : cat3
  Auto reservice times/limit       : 0/10
  Digitmap group                 : DigitmapGroup1
  Incoming called number change plan : NumberChange1
  Outgoing caller number change plan : NumberChange2
  CLI number                    : 12345678
  CLI mode                      : match
  Bear capability                : speech
  Called number type             : national
  International prefix flag       : enable
  National prefix flag           : enable
  DSP-para-template              : -
  ------------------------------------------------------------------------------
huawei(config-esl-user)#sipbrauser extend-telno add 0/4/0 +86-755-28780000
huawei(config-esl-user)#display sipbrauser extend-telno 0/4/0
  ------------------------------------------------------------
    Index          Extend-telno
   ------------------------------------------------------------
    1             +86-755-28780000
    2             28780000
    3             0755-28780000
   ------------------------------------------------------------
```

(Optional) Configuring the System Parameters

This topic describes how to configure the system parameters including the overseas version flag and message waiting indication (MWI) mode according to the local standards to ensure that the response of the user terminal complies with the local standards.

Procedure

Run the **system parameters** command to configure the system parameters.

Step 1 Run the **display system parameters** command to check whether the system parameters are the same as those in the data plan.

----End

Example

To configure the overseas version flag (system parameter 1) to Hong Kong (parameter value 1), do as follows:

```
huawei(config)#system parameters 1 1
huawei(config)#display system parameters 1
  ---------------------------------------------------------------------------
  Parameter name index: 1    Parameter value: 1
  Mean: Overseas version flag, 0:China, 1:HongKong, 2:Brazil, 3:Egypt, 4:
Singapore, 5:Thailand, 6:France, 7:Britain MSFUK, 8:Britain ETSI, 9:Bulgaria,
10:Reserved, 11:Austria, 12:Hungary, 13:Poland
  ---------------------------------------------------------------------------
```

(Optional) Configuring the Overseas Parameters

By default, the overseas parameters are configured according to the Chinese standards. In the actual service configuration, the attributes such as the upper and lower thresholds of the flash-hooking duration can be configured according to the local standards to ensure that the response of the user terminal complies with the local standards.

Procedure

Run the **oversea parameters** command to configure the overseas parameters.

Step 1 Run the **display oversea parameters** command to check whether the overseas parameters are the same as those in the data plan.

----End

Example

To set the upper flash-hooking threshold (overseas feature parameter 0) to 800 ms (in compliance with the Hong Kong standard) and the lower flash-hooking threshold (overseas feature parameter 1) to 100 ms (in compliance with the Hong Kong standard), do as follows:

```
huawei(config)#oversea parameters 0 800
huawei(config)#oversea parameters 1 100
huawei(config)#display oversea parameters
```

```
{ <cr>|name<U><0,99> }:

 Command:
       display oversea parameters
 ------------------------------------------------------------------------------
 Parameter name index: 0    Parameter value: 800
 Mean: Hooking upper threshold(ms), reference: China:350, HongKong:800
 ------------------------------------------------------------------------------
 Parameter name index: 1    Parameter value: 100
 Mean: Hooking lower threshold(ms), reference: China:100, HongKong:100
 ------------------------------------------------------------------------------
 Parameter name index: 2    Parameter value: 0
 Mean: Flag of applying PARKED LINE FEED or not when user port is locked, 0:not
apply, 1:apply
 ------------------------------------------------------------------------------
 Parameter name index: 3    Parameter value: 0
 Mean: The detect time of flash upper limit to onhook, default value: 0ms
 ------------------------------------------------------------------------------
 Parameter name index: 4    Parameter value: 0
 Mean: DTMF Detector Tuning, 0:Q.24, 1:Russia
------------------------------------------------------------------------------
 Parameter name index: 5    Parameter value: 0
 Mean: Flag of changing TEI value for software switch, 0:no change, 1:set TEI
value to 1, 2:set TEI value to 0. Default: 0
 ------------------------------------------------------------------------------
```

(Optional) Configuring the Attributes of the ISDN BRA Port

This topic describes how to configure the attributes of an ISDN BRA port to ensure that the ISDN BRA port can meet the actual application requirements. You can configure the auto-deactivation status, remote power supply status, UNI fault alarming function, and working mode of the port.

Default Configuration

Table 23-31 lists the default values of the attributes of an ISDN BRA port. When configuring the attributes, you can change the values according to the service requirements.

Table 23-31 Default values of the attributes of an ISDN BRA port

Parameter	Default Setting
Autodeactive	Disable
Autodeactive-delay	30s
Activemode	unstable-active
Remotepower	Disable
Unialarm	Disable
Workmode	p2mp

Procedure

In global config mode, run the **braport** command to enter braport mode.

Step 1 Run the **braport attribute set** command to configure the attributes such as the working mode, auto-deactivation status, and remote power supply status of the port.

If an ISDN BRA port needs to be connected to multiple terminal users, configure the working mode of the port to p2mp. If an ISDN BRA port needs to be connected to only one terminal user, configure the working mode of the port to p2p.

For detailed description of the **braport attribute set** command, see the parameter description in **braport attribute set**.

----End

Example

Assume that the working mode is p2mp, the activation mode is stable, and the auto-deactivation function is disabled. To configure such attributes of ISDN BRA port 0/4/0, do as follows:

```
huawei(config)#braport
huawei(config-braport)#braport attribute set 0/4/0 workmode p2mp activemode
stable-active
huawei(config-braport)#display braport attribute
{ frameid/slotid/portid<S><Length 1-15>|frameid/slotid<S><Length 1-15> }:0/4/0

  Command:
         display braport attribute 0/4/0
-------------------------------------------------------------------------------
F  /S /P  Remotepower Workmode Autodeactive Deactivedelay Activemode Unialarm
-------------------------------------------------------------------------------
0/4/0  disable    p2mp    disable     30          stable    disable
-------------------------------------------------------------------------------
```

23.19 Configuring the VoIP ISDN PRA Service (SIP-based)

After configuring the SIP interface, you can add VoIP ISDN users on this interface to implement the VoIP ISDN service.

Prerequisites

According to the actual network, a route from the Access node to the IMS must be configured to ensure that the Access node communicates with the IMS normally.

Context

- The voice over IP (VoIP) service uses the IP packet switched network for transmission after the traditional analog voice signals are compressed and packetized. This service

lowers the cost of the voice service. For the detailed description of the VoIP service, see 23 Voice Feature in the *Feature Description*.

- Defined by the International Telegraph and Telephone Consultative Committee (CCITT), the ISDN is a communication network evolved from the Integrated Digital Network (IDN). The ISDN service provides the E2E digital connection and supports multiple types of voice and non-voice telecom services. On the ISDN network, users can access the network through the following two interfaces: (The ISDN technology provides two types of user-network interfaces based on the user-network interface reference model.)
  - ISDN basic rate interface (BRI), which supports a rate of 144 kbit/s and provides two B channels (for carrying services) and one D channel (for transmitting call control signaling and maintenance and management signaling). The rate of B channels is 64 kbit/s and the rate of D channel is 16 kbit/s. The service carried on the ISDN BRI is the ISDN basic rate access (BRA) service.
  - ISDN primary rate interface (PRI), which supports a rate of 2.048 Mbit/s and provides 30 B channels and one D channel. The rates of the B channel and D channel are both 64 kbit/s. The service carried on the ISDN PRI is the ISDN primary rate access (PRA) service.

The Access node can also function as a voice over IP gateway (VGW) in the IMS architecture. In the downstream direction, it is connected to the ISDN PRA users; in the upstream direction, it is connected to the IMS system through the SIP interface, providing the VoIP ISDN PRA service by working with the IMS core.

The functions and applications of the SIP interface are the same as the functions and applications of the MG interface.

Data preparation

Table 23-32 provides the data plan for configuring the VoIP ISDN PRA service.

Table 23-32 Data plan for configuring the VoIP ISDN PRA service when the SIP protocol is used

Item			Remarks
SIP interface data	Media and signaling parameters	Media and signaling upstream VLAN	It is used for the upstream VLAN of the VoIP service to be configured. **NOTICE** Note that the media and the signaling can use the same VLAN or different VLANs, depending on the negotiation with the upstream device.
		Signaling upstream port	Uplink port for configuring the SIP signaling.
		Media IP address and signaling IP addresses	These IP addresses must be selected from the media and signaling IP address pools. The media and signaling IP address pools consists of all the IP addresses of the L3 interface of the media and signaling upstream VLAN.
		Default IP address of the	Next hop address from the Access node to the IMS core network device.

Item			Remarks
		MG	**NOTICE** If the media IP address and the signaling IP address are different and the media and the signaling are transmitted upstream through different gateways, ensure that they correspond to correct gateways. Otherwise, the call service may fail.
	Attributes of the SIP interface **NOTE** Parameters listed here are mandatory, which means that the SIP interface cannot be enabled if these parameters are not configured.	SIP interface ID	It is used for the VoIP service to be configured.
		Signaling port ID of the SIP interface	The value range is 5000-5999. The protocol defines the port ID as 5060.
		IP address of the active IMS core network device to which the SIP interface belongs	When dual homing is not configured, parameters of only one IMS core network device are required. When dual homing is required, the IP address and the port ID of the standby IMS core network device must be configured.
		Port ID of the active IMS core network device to which the SIP interface belongs	
		Transmission mode of the SIP interface	The transmission mode is selected according to the requirements of the IMS core network device. Generally, UDP is used.
		Home domain of the SIP interface	It corresponds to parameter **home-domain** in the MG interface attributes.
		Index of the profile used by the SIP interface	It corresponds to parameter **Profile-index** in the MG interface attributes.
		IP address obtaining mode of the proxy server	• In the IP mode, the IP address and the port ID of the active proxy server must be configured. • In the DNS-A or DNS-SRV mode, the domain name of the active proxy server must be configured.
ISDN PRA user data (The data configuration must be consistent with the data configuration	Slot that houses the E1 service board.		-
	Configuring the ISDN PRA User Data	Phone number	The phone number that the IMS core network device allocates to the user must be configured.
		User priority	According to the service requirements, user priorities must be specified,

Item			Remarks
on the IMS.)			including: • cat1: government1 (category 1 government users) • cat2: government2 (category 2 government users) • cat3: common (default priority, namely, common users)
	(Optional) Configuring the Centrex		The out-centrex prefix and out-centrex attributes of a centrex need to be configured according to local standards.
	(Optional) Configuring the System Parameters		The system parameters including the oversea version flag and message waiting indication (MWI) mode need to be configured according to local standards to ensure that the response of the user terminal meets local standards.
	(Optional) Configuring the Overseas Parameters		The attributes such as the upper and lower thresholds of the flash-hooking duration need to be configured according to local standards to ensure that the response of the user terminal meets local standards.
	Configuring the Attributes of the E1 Port		Board access mode, port mode, line coding mode, and port impedance need to be configured, and need not be modified if there is no special requirement.

23.19.1 Configuring the SIP Interface

As an interface used for the intercommunication between the MA5600T/MA5603T/MA5608T and the MGC, the interface is vital to the SIP-based VoIP service. Therefore, to implement the VoIP service, the SIP interface must be configured and must be in the normal state.

Procedure

Configuring the Upstream VLAN Interface

This topic describes how to specify the upstream VLAN interface for the media stream and the signaling flow, and how to configure the IP addresses of the Layer 3 interface. These IP addresses are the sources of the media and signaling IP address pools.

Context

Multiple IP addresses can be configured for the same VLAN Layer 3 interface. Only one IP address functions as the primary address, and other IP addresses function as the secondary addresses.

Procedure

Run the **vlan** command to add an upstream VLAN for the media stream and the signaling flow.

Step 1 Run the **port vlan** command to add the upstream ports to the VLAN.

Step 2 Configure the IP addresses of the VLAN Layer 3 interface.

1. Run the **interface vlanif** command to enter the Layer 3 interface of the upstream VLAN for the media stream and the signaling flow.
2. Run the **ip address** command to configure the IP addresses of the Layer 3 interface.

Step 3 Run the **display interface vlanif** command to check whether the IP addresses of the Layer 3 interface are the same as those in the data plan.

----End

Example

Assume that the media stream and the signaling stream are transmitted upstream through upstream port 0/19/0. To create media and signaling upstream VLAN 3 and configure IP addresses 10.13.4.116/16 and 10.13.4.117/16 of the Layer 3 interfaces for the media IP address pool and the signaling IP address pool respectively, do as follows:

```
huawei(config)#vlan 3 standard
huawei(config)#port vlan 3 0/19 0
huawei(config)#interface vlanif 3
huawei(config-if-vlanif3)#ip address 10.13.4.116 16
huawei(config-if-vlanif3)#ip address 10.13.4.117 16 sub
huawei(config-if-vlanif3)#quit
huawei(config)#display interface vlanif 3
vlanif3 current state : UP
Line protocol current state : UP
Description : HUAWEI, SmartAX Series, vlanif3 Interface
The Maximum Transmit Unit is 1500 bytes
Internet Address is 10.13.4.116/16
Internet Address is 10.13.4.117/16 Secondary
IP Sending Frames' Format is PKTFMT_ETHNT_2, Hardware address is 00e0-fc32-1118
```

Configuring the Media and Signaling IP Address Pools

The media IP address pool defines all media IP addresses that can be used by the AG, and the signaling IP address pool defines all signaling IP addresses that can be used by the AG.

Prerequisites

The IP address of the Layer 3 interface of the media and signaling upstream VLAN must be configured. For details about how to configure the IP address, see Configuring the Upstream VLAN Interface.

Context

- The media IP address and the signaling IP address for the MG interface must be selected from the IP address pools configured here.
- The signaling IP address pool is used to store the IP addresses of the MG interfaces, and the media IP address pool is used to store the IP addresses of the media streams controlled by the signaling.
- The media IP address pool and the signaling IP address pool can be the same. Similarly, the media IP address and the signaling IP address can be the same.

NOTICE

The MGCP interface on the MA5600T/MA5603T/MA5608T does not support the isolation of media streams and signaling flows. Therefore, when the MGCP protocol is used, the media IP address pool and the signaling IP address pool must be the same. When the H.248 or SIP protocol is used, the media IP address pool and the signaling IP address pool can be the same or different.

Procedure

Run the **voip** command to enter the VoIP mode.

Step 1 Configure the media IP address pool.

1. Run the **ip address media** command to add the media IP address to the media IP address pool.

   The media IP address needs to be selected from the IP addresses of the Layer 3 interface of the media and signaling upstream VLAN.
2. Run the **display ip address media** command to check whether the media IP address pool is the same as that in the data plan.

Step 2 Configure the signaling IP address pool.

1. Run the **ip address signaling** command to add the signaling IP address to the signaling IP address pool.

   The signaling IP address needs to be selected from the IP addresses of the Layer 3 interface of the media and signaling upstream VLAN.
2. Run the **display ip address signaling** command to check whether the signaling IP address pool is the same as that in the data plan.

----End

Example

To add IP addresses 10.13.4.116/16 and 10.13.4.117/16 of Layer 3 interfaces of the media and signaling upstream VLAN to the media IP address pool and the signaling IP address pool respectively, do as follows:

```
huawei(config)#voip
huawei(config-voip)#ip address media 10.13.4.116 10.13.0.1
huawei(config-voip)#ip address media 10.13.4.117 10.13.0.1
huawei(config-voip)#ip address signaling 10.13.4.116
```

```
huawei(config-voip)#ip address signaling 10.13.4.117
huawei(config-voip)#display ip address media
  Media:
  IP Address...........: 10.13.4.116
  Subnet Mask..........: 255.255.0.0
  Gateway..............: 10.13.0.1
  MAC Address..........: 00-E0-FC-AF-91-33

  IP Address...........: 10.13.4.117
  Subnet Mask..........: 255.255.0.0
  Gateway..............: 10.13.0.1
  MAC Address..........: 00-E0-FC-AF-91-33
huawei(config-voip)#display ip address signaling
  Signaling:
  IP Address...........: 10.13.4.116
  Subnet Mask..........: 255.255.0.0
  MAC Address..........: 00-E0-FC-AF-91-33

  IP Address...........: 10.13.4.117
  Subnet Mask..........: 255.255.0.0
  MAC Address..........: 00-E0-FC-AF-91-33
```

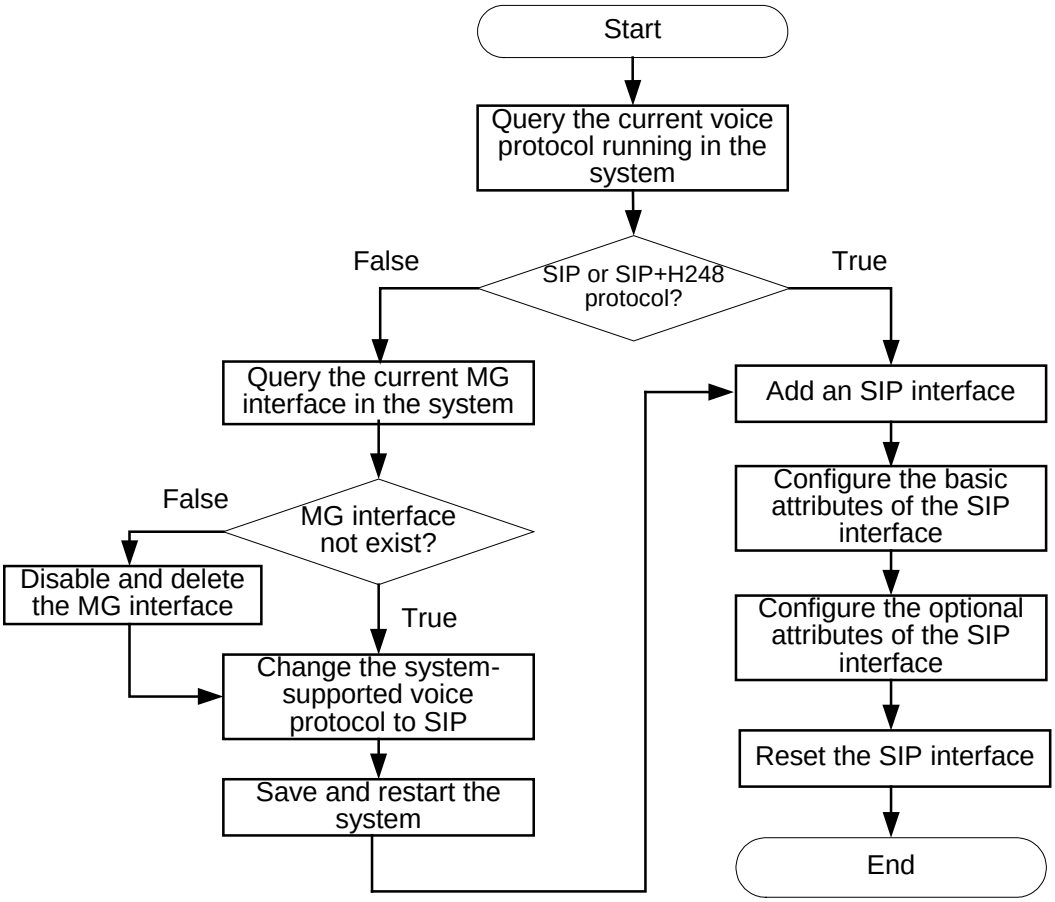
Adding an SIP Interface

The MA5600T/MA5603T/MA5608T exchanges the signaling and protocol packets with the information management system (IMS) through an SIP interface. To ensure uninterrupted signaling and protocol packet exchange, ensure that the status of the SIP interface is in a normal state after you add it.

Context

- One MA5600T/MA5603T/MA5608T supports up to eight SIP interfaces. Each SIP interface can be configured with the interface attributes separately.
- The SIP attributes configured for an SIP interface take effect on this interface only.
- The attributes of an SIP interface, including the signaling IP address, media IP address, transport-layer protocol, port ID of the transport-layer protocol, IP address (or domain name) of the proxy server, port ID of the proxy server, home domain name, profile index, are mandatory. In addition, the attributes of an SIP interface must be consistent with those configured on the IMS side so that the status of the SIP interface is normal.

Configuration Flowchart



Procedure

Query the current voice protocol running in the system.

Run the **display protocol support** command to query the voice protocol that is currently supported by the system.

- If the system voice protocol is the SIP protocol, go to Step 6.
- If the system voice protocol is not the SIP protocol, go to Step 2.

Step 1 Query the current MG interface in the system.

Run the **display if-mgcp all** or **display if-h248 all** command to query whether an MGCP interface or an H.248 interface currently exists in the system.

- If there is no such an MG interface, go to Step 4.
- If there is such an MG interface, go to Step 3.

Step 2 Disable and delete the MG interface.

1. Delete the user data of this MG interface, and then run the **shutdown(mgcp)** or **shutdown(h248)** command to disable the MG interface according to the protocol type of the interface.

NOTICE

This operation interrupts all the ongoing services carried on the MG interface. Hence, exercise caution when performing this operation. Before performing this operation, you must check whether the operation is allowed.

2. Run the **undo interface mgcp** or **undo interface h248** command to delete the MG interface.

Step 3 Change the system-supported voice protocol to SIP.

Run the **protocol support** command to change the system-supported voice protocol to SIP.

Step 4 Save the configuration data and restart the system.

Save the configuration data by running the **save** command. Then run the **reboot** command to restart the system to make the new configuration data take effect.

Step 5 Run the **interface sip** command to add an SIP interface.

Step 6 Run the **if-sip attribute basic** command to configure the basic attributes of the SIP interface.

NOTE

- Between **proxy IP** and **proxy domain**, which are basic attributes, at least one attribute must be configured. If both attributes are configured, the system determines which attribute is used according to the configured address mode of the proxy server.
- The profile index must be configured.

Step 7 Run the **if-sip attribute optional** command to configure the optional attributes of the SIP interface.

The optional attributes include some new service types supported by the SIP interface, such as the conference call, message waiting indicator, UA-profile subscription, and REG-state subscription, and also include the SIP proxy configuration mode. The optional attributes require the IMS-side support, and must be consistent with those on the IMA-side.

After the configuration is completed, run the **display if-sip attribute config** command to query the attributes of the SIP interface.

Step 8 Reset the SIP interface.

Run the **reset** command to reset the SIP interface for the new configuration data to take effect. Otherwise, the configuration data does not take effect but is only stored in the database.

After the SIP interface is reset successfully, run the **display if-sip attribute running** command to query the running status of the SIP interface. If the active (or standby) proxy is **up**, the SIP interface is normal.

----End

Example

Assume that: the media IP address is 10.10.10.13, signaling IP address is 10.10.10.13, transmission protocol is UDP, port ID is 5000, active proxy server IP address 1 is 10.10.10.14, port ID of the active proxy server is 5060, active proxy server domain name is proxy.domain, standby proxy server IP address 1 is 10.10.10.15, port ID of the standby proxy server is 5060,

home domain name is sip.huawei.com, and profile index is 1. To configure the attributes of SIP interface 0, do as follows:

```
huawei(config)#interface sip 0
huawei(config-if-sip-0)#if-sip attribute basic media-ip 10.10.10.13 signal-ip 10
.10.10.13 signal-port 5000 transfer udp primary-proxy-ip1 10.10.10.14 primary-pr
oxy-port 5060 primary-proxy-domain proxy.domain secondary-proxy-ip1 10.10.10.15
secondary-proxy-port 5060 home-domain sip.huawei.com sipprofile-index 1
huawei(config-if-sip-0)#reset
  Are you sure to reset the SIP interface?(y/n)[n]:y
huawei(config-if-sip-0)#
  Resetting SIP interface 0 succeeded
huawei(config-if-sip-0)#display if-sip attribute running
  ------------------------------------------------------------
  ...//The rest information in response to this command is omitted.
  Primary Proxy State               up   //Indicates that the SIP interface is in the
normal state.
  Secondary Proxy State              down
  ...
  ------------------------------------------------------------
```

(Optional)Configuring the Software Parameters of an SIP Interface

The software parameters of a SIP interface mainly define certain common service attributes of the SIP interface. After the software parameter configuration, the parameters take effect immediately and are valid only to the SIP interface. Skip this topic if the system default configuration meets user requirements.

Prerequisites

The MG interface has been configured. For details about how to configure the MG interface, see 23.19.1 Configuring the SIP Interface.

Context

The details about the software parameters that can be configured of a SIP interface that supports SIP, see section **Usage Guidelines** in **mg-software parameter**. The other parameters are reserved in the system.

Table 23-33 lists parameters that are usually configured to a non-default value. The other parameters are not required.

Table 23-33 Software parameters usually configured of a SIP interface

Parameter	Description	Default Setting
2	Indicates whether the standalone mode is supported.	Numeral type. Range: 0-1. • 0: indicates that the standalone function is not supported. • 1: indicates that the standalone function is

Parameter	Description	Default Setting
		supported. Default: 0 This parameter is usually set to **1**.
8	Indicates whether the heartbeat message of the MG is disabled.	Numeral type. Range: 0-1. • 0: the heartbeat message of the MG is disabled • 1: the heartbeat message of the MG is enabled Default value: 0.

Procedure

Enter the SIP interface mode.

In global config mode, run the **interface sip** command to enter the SIP interface mode.

Step 1 Configure software parameters.

Run the **mg-software parameter** command the software parameters required in the data plan.

Step 2 Check whether the software parameters are the same as those in the data plan.

Run the **display mg-software parameter** command to check whether the software parameters of the MG interface are the same as those in the data plan.

----End

Example

To configure software parameter 2 of the SIP interface 0 to 1 so that the SIP interface supports standalone, do as follows:

```
huawei(config)#interface sip 0
huawei(config-if-sip-0)#mg-software parameter 2 1
huawei(config-if-sip-0)#display mg-software parameter 2
  -------------------------------------------------
  MGID:0         para index:2   value:1
  -------------------------------------------------
 APPENDIX:
  -------------------------------------------------
 Parameter Index:  Interface software parameter name:
    2 : SAL Support
        0: No
        1: Yes
```

23.19.2 Configuring the VoIP ISDN PRA User

This topic describes how to configure the VoIP ISDN PRA user. After the MG interface is configured, you can add the VoIP ISDN PRA user on this interface to implement the VoIP ISDN PRA service.

Configuring the Attributes of the E1 Port

This topic describes how to configure the attributes of the E1 port to ensure that the ISDN PRA port meets the actual application requirements.

Context

You can configure the impedance, line coding mode, and working mode of the E1 port.

Default Configuration

Table 23-34 lists the default values of the E1 port. When configuring the attributes of the E1 port, you need to modify the values according to the service requirements.

Table 23-34 Default values of the E1 port

Parameter	Default Setting
Port impedance	E1 mode: 75 ohm
Line coding mode	E1 mode: HDB3
CRC4	Enable
The mode for digital section access	Digital
Signaling type	CCS

Procedure

In global config mode, run the **interface edt** command to enter the EDT mode.

Step 1 (Optional) Run the **e1port impedance** command to configure the impedance of an E1 port.

Step 2 (Optional; perform this step when you need to modify the line coding mode of the port) Run the **e1port line-code** command to configure the line coding mode of the E1 port.

NOTE
In E1 mode, the system supports two line coding modes, namely, HDB3 and AMI.

Step 3 (Optional) Run the **e1port crc4** command to configure the CRC4 function of an E1 port.

Step 4 (Optional) Run the **e1port attribute set** command to configure the digital section access mode of an E1 port.

Step 5 (Optional) Run the **e1port signal** command to configure the signaling type of an E1 port.

----End

Example

Assume that the E1 ports on ISDN PRA board work in HDB3 line encoding mode, the CRC4 function is enable, To configure such E1 ports, do as follows:

```
huawei(config)#interface edt 0/1
huawei(config-if-edt-0/1)#e1port line-code 1 HDB3
huawei(config-if-edt-0/1)#display e1port line-code 1
  ---------------------
  F/S/P   linecode
  ---------------------
  0/1/1   HDB3
  ---------------------
huawei(config-if-edt-0/1)#e1port crc4 1 enable
huawei(config-if-edt-0/1)#display e1port attribute 1
  ------------------------------------------------------------
   F/S/P    Signaltype    CRC4    Impedance    Accessmode
  ------------------------------------------------------------
   0/1/1    CCS          Enable  75          Digital
  ------------------------------------------------------------
```

Configuring the ISDN PRA User Data

This topic describes how to configure the ISDN PRA user data on the SIP interface (the data must be the same as the corresponding data on the IMS) so that the ISDN PRA user can access the network to use the ISDN PRA service.

Prerequisites

The ISDN PRA service board must be inserted into the planned slot correctly and the board must be in the normal state.

Default Configuration

Table 23-35 lists the default settings of the attributes of the ISDN PRA user. When configuring the attributes of these attributes, you can modify the values according to the service requirements.

Table 23-35 Default settings of the attributes of the ISDN PRA user

Parameter	Default Settings
Priority of the ISDN PRA user	cat3 (common users)
Flag of reporting the UNI fault of the ISDN PRA user	disable
Sub-channel active mask of the ISDN PRA user	255.255.255.255
Threshold for the number of auto recoveries from deterioration faults	20
The matching scheme of an outgoing calling number	match

Parameter	Default Settings
The voice bear capability of a port	speech+3.1k-audio
The type of the called number sent to the ISDN terminal when an incoming call is initiated	unknown
The international prefix flag	disable
The national prefix flag	disable
The change plan of an incoming called number used by the user on the port	-
The change plan of an outgoing calling number used by the user on the port	-
The calling number	-
The digitmap group used by the user	-

Procedure

In global config mode, run the **esl user** command to enter the ESL user mode.

Step 1 Run the **sipprauser add** command to add an ISDN PRA user.

Step 2 Run the **display sipprauser** command to check whether the ISDN PRA user data is the same as the data plan.

Step 3 (Perform this step when you need to modify the attributes of an ISDN PRA user.) Run the **sipprauser attribute set** command to configure the attributes of the ISDN PRA user.

Step 4 (Perform this step only after you modify the attributes of the ISDN PRA user.) Run the **display sipprauser attribute** command to query whether the configured attributes of the ISDN PRA user are the same as the data plan.

Step 5 (Perform this step only when you need to configure an extended phone number for the ISDN PRA user.) Run the **sipprauser extend-telno add** command to add multiple phone numbers or a phone number containing non-digit characters for the ISDN PRA user.

Step 6 (Perform this step only after you configure the extended phone number for an ISDN PRA user.) Run the **display sipprauser extend-telno** command to query whether the configured extended phone number of the ISDN PRA user is the same as the data plan.

----End

Example

Assume that:

- SIP interface ID: 0
- Phone number: 28780000
- Call priority: cat3
- UNI fault report flag: disable

- Number of auto recoveries from deterioration faults: 10
- Sub-channel active mask: 255.255.1.3
- Voice bear capability: speech
- Type of the called number: national
- International prefix flag: enable
- national prefix flag: enable
- Digitmap group: DigitmapGroup1
- Change plan of an incoming called number: NumberChange1
- Change plan of an outgoing calling number: NumberChange2
- Calling number is 12345678
- Extended phone number: +86-755-28780000

To add such an ISDN PRA user connected to the 0/1/0 port, do as follows:

```
huawei(config)#esl-user
huawei(config-esl-user)#sipprauser add 0/1/0 0 telno 28780000
huawei(config-esl-user)#display sipprauser 0/1/0
{ <cr>|endframeid/slotid/portid<S><Length 1-15> }:

  Command:
         display sipprauser 0/1/0
  ---------------------------------------
  F  /S /P    MGID    TelNo
  ---------------------------------------
  0/1/0   0       28780000
  ---------------------------------------
huawei(config-esl-user)#sipprauser attribute set 0/1/0 priority cat3 unireport d
isable auto-resume-limit 10 activemask 255.255.255.3 bc speech called-num-type n
ational international-prefix-flag enable national-prefix-flag enable digitmap-gr
oup DigitmapGroup1 incomingcall-numberchange NumberChange1 outgoingcall-numberch
ange NumberChange2 cli-number 12345678
huawei(config-esl-user)#display sipprauser attribute 0/1/0
{ <cr>|endframeid/slotid/portid<S><Length 1-15> }:

  Command:
         display sipprauser attribute 0/1/0
  ------------------------------------------------------------------------
  F  /S /P                       : 0/1/0
  UNIreport                      : disable
  Prior                          : cat3
  Mask of sub channel              : 255.255.255.3
  Auto reservice times/limit       : 0/10
  Digitmap group                 : DigitmapGroup1
  Incoming called number change plan : NumberChange1
  Outgoing caller number change plan : NumberChange2
  CLI number                     : 12345678
  CLI mode                       : match
  Bear capability                : speech+3.1k-audio
  Called number type             : unknown
  International prefix flag        : disable
  National prefix flag           : disable
  ------------------------------------------------------------------------
huawei(config-esl-user)#sipprauser extend-telno add 0/1/0 +86-755-28780000
```

```
huawei(config-esl-user)#display sipprauser extend-telno 0/1/0
  ------------------------------------------------------------
    Index          Extend-telno
  ------------------------------------------------------------
    1              +86-755-28780000
    2              28780000
    3              0755-28780000
  ------------------------------------------------------------
```

(Optional) Configuring the Centrex

Centrex refers to a virtual user group. The MA5600T/MA5603T/MA5608T supports the following functions: Members in a centrex can call each other by dialing short numbers, and members in a centrex can call the members outside of the centrex by dialing centrex prefix + the complete phone number. Generally, a softswitch issues centrex parameters. If a softswitch does not issue centrex parameters, use the parameters preset on the MA5600T/MA5603T/MA5608T. These parameter values must be the same as these on the softswitch.

Context

- Centrex prefix: When attempting to call a user in another centrex group, a user must dial the centrex prefix before dialing the called number. A centrex prefix contains 0 to 9 digits.
- The function that the members in a centrex can call each other by dialing short numbers need not be configured on the MA5600T/MA5603T/MA5608T through the command line interface (CLI).
- The function that the members in a centrex can call the members outside of the centrex by dialing centrex prefix + the complete phone number can be supported only when the SIP protocol is used.
- The centrex attribute of a centrex can be direct centrex or two-stage centrex. The similarity and difference are as follows:
  - Similarity: When the members in a centrex need to call the members outside of the centrex, they must dial the centrex prefix.
  - Difference: If the centrex attribute is set to two-stage centrex, the members in a centrex can hear the dial tone again after dialing the centrex prefix. If the centrex attribute is set to direct centrex, no out-group dial tone is played.

Procedure

Configure the centrex call function for a centrex group.

The MA5600T/MA5603T/MA5608T supports the configuration of the centrex prefix through one of two methods. In method 1, configure the centrex prefix and centrex attributes for a single user in ESL user mode. In method 2, configure the centrex prefix and centrex attributes for all the users in global config mode. When both methods are used, method 1 takes effect.

- Use method 1:

1. In ESL user mode, run the **sippstnuser servicedata parameter set** command to configure the centrex prefix and centrex attributes of a centrex group.
2. Run the **display sippstnuser servicedata** command to check whether the centrex parameter settings of a centrex are the same as the data plan.

- Use method 2:

NOTE

If method 2 is used, the MA5600T/MA5603T/MA5608T uses the centrex digitmap to match the centrex prefix, and uses the call digitmap, or the normal digitmap, to match the phone number dialed by a user.

1. In global config mode, run the **local-digitmap add** command to configure a direct centrex digitmap or a two-stage centrex digitmap.

NOTE

- The system does not support the adding of a direct centrex digitmap and a two-stage centrex digitmap at the same time. Add a digitmap based on site requirements.
- When you add a direct centrex digitmap, the system centrex attribute is direct centrex. When you add a two-stage centrex digitmap, the system centrex attribute is two-stage centrex.

2. Run the **display local-digitmap** command to check whether the local digitmap is the same as the data plan.

Step 1 Check whether the value of the sipprofile control point 148 is the same as the data plan.

Run the **display sipprofile syspara detail** command to check whether the value of control point 148 is the same as the data plan. If they are different, run the **sipprofile modify** command in SIP mode to change the control point value.

NOTE

- The sipprofile control point 148 can be set to 0 or 1. When it is set to 0, a phone number does not contain a centrex prefix; when it is set to 1, a phone number contain a centrex prefix. By default, it is 1.
- Run the **if-sip attribute basic** *sipprofile-index 0* command to specify a user-defined profile for the current SIP interface, and run the **sipprofile modify** command to change the value for the sipprofile control point 148.

----End

MA5600T/MA5603T/MA5608T

Assume that the centrex prefix of the MA5600T/MA5603T/MA5608T user with phone number 88627792 is 8100, the centrex attribute is two-stage centrex, and the control point of the Sipprofile uses the default value.

To configure the centrex call function for such a user by using method 1, do as follows:

```
huawei(config)#esl user
huawei(config-esl-user)#sippstnuser servicedata parameter set 0/2/1 telno 88627792
centrexprefix 8100 centrexflag dialsecondary
huawei(config-esl-user)#display sippstnuser servicedata 0/2/1 telno 88627792
  ----------------------------------------
  F /S /P              : 0/2/1
  telno                : 88627792
  centrexno            : -
  centrexprefix        : 8100
  centrexflag          : dialsecondary
  mwimode              : deferred
  hottime(s)           : 100
  hotlinenum           : -
  dialtone             : normal
  cfbnum               : -
  cfnrnum              : -
  cfunum               : -
  cfnrtime(s)          : 100
```

```
  displayname           : -
  permanent-hold-mode    : norecall
  permanent-hold-time(s)  : 20
  ------------------------------------------------
```

To configure the centrex call function for such a user by using method 2, and plan the digitmap body to (8100) and the digitmap name to **huawei1** for the two-stage centrex digitmap according to the centrex prefix, do as follows:

```
huawei(config)#local-digitmap add huawei1 second-centrex (8100)
huawei(config)#display local-digitmap all
  ----------------------------------------
  Name   : huawei1
  Type   : second-centrex
  Body   : (8100)
  Protocol: sip
  ----------------------------------------
```

(Optional) Configuring the System Parameters

This topic describes how to configure the system parameters including the overseas version flag and message waiting indication (MWI) mode according to the local standards to ensure that the response of the user terminal complies with the local standards.

Procedure

Run the **system parameters** command to configure the system parameters.

Step 1 Run the **display system parameters** command to check whether the system parameters are the same as those in the data plan.

----End

Example

To configure the overseas version flag (system parameter 1) to Hong Kong (parameter value 1), do as follows:

```
huawei(config)#system parameters 1 1
huawei(config)#display system parameters 1
  ------------------------------------------------------------------------------
  Parameter name index: 1    Parameter value: 1
  Mean: Overseas version flag, 0:China, 1:HongKong, 2:Brazil, 3:Egypt, 4:
Singapore, 5:Thailand, 6:France, 7:Britain MSFUK, 8:Britain ETSI, 9:Bulgaria,
10:Reserved, 11:Austria, 12:Hungary, 13:Poland
  ------------------------------------------------------------------------------
```

(Optional) Configuring the Overseas Parameters

Address Resolution Protocol (ARP) probe enables faster protection switching by detecting status of end-to-end links. ARP detection can be configured for a network scenario in which a link protection group is configured for the upstream Ethernet ports on the access device, and

there are other types of devices deployed, such as switches and transmission devices, between the access device and the aggregation devices.

Prerequisites

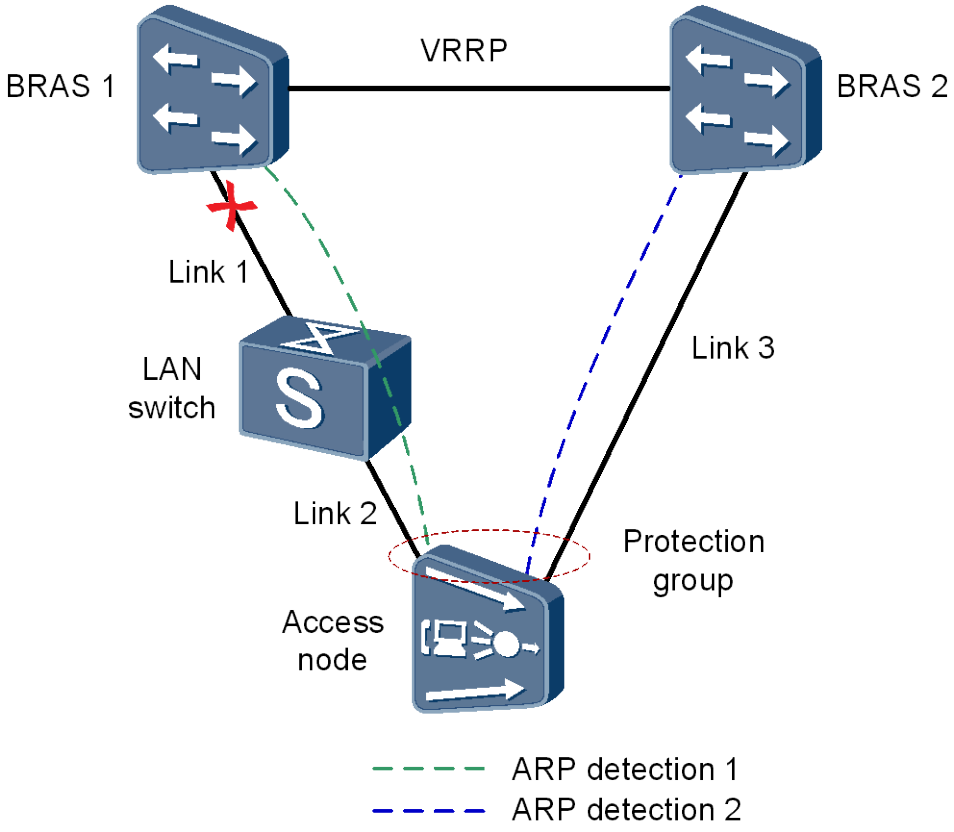
The access device provides upstream ports through the upstream boards. ARP detection is not supported if the access device provides upstream ports through the control board.

A Timedelay protection group is configured on the upstream boards. For configuration details, see 19.4.5 Configuring an Ethernet Port Protection Group.

Context

Figure 23-132 shows an Ethernet port protection example in which the access device (MA5600T/MA5603T/MA5608T) is dual-homed to BRASs.

Figure 23-132 Ethernet port protection with the access device dual-homed to BRASs



The MA5600T/MA5603T/MA5608T is dual-homed to BRAS 1 and BRAS 2. The working links are Link1 and Link2, and the protection link is Link3. If Link1 is faulty and Link2 is normal, the MA5600T/MA5603T/MA5608T can switch services to Link3 by using ARP detection even though the upstream port on the MA5600T/MA5603T/MA5608T is functioning properly. This ensures uninterrupted service transmission.

Procedure

Configure VLAN and Layer 3 port IP address for ARP detection.

1. Create a VLAN.

   Run the **vlan** command to create a smart VLAN for ARP detection.

2. Add an upstream port to the VLAN.

   Run the **port vlan** command to add the working upstream Ethernet port in the protection group to the VLAN. The protection port in the protection group cannot be added to the VLAN.

3. Create a Layer 3 interface for the VLAN.

   Run the **interface vlanif** command to create a Layer 3 interface for the VLAN and enter the VLAN interface mode.

4. Configure an IP address for the Layer 3 interface in the VLAN.

   Run the **ip address** command to configure an IP address for the Layer 3 interface in the VLAN. Ensure that this IP address is in the same subnet as the IP address of the remote device.

Step 1 Configure ARP detection for the working port in the protection group.

1. Configure ARP detection for the working port in the protection group.

   Run the **arp-detect** command to configure ARP detection for the working port in the protection group.

2. (Optional) Configure the times for sending ARP detection packets.

   Run the **detect-multiplier** command to configure the times for sending ARP packets. If the remote device does not respond to the ARP detection packets sent by the local device (here, the access device) for the specified times, the local device considers ARP detection has timed out.

   The waiting time for ARP detection is derived from the following formula: Waiting time = Interval for sending ARP request packets x Times for sending ARP packets. The ARP detection waiting time is also the time taken for triggering a protection switching. The minimum waiting time is 3s (1s x 3). Because the interconnected devices have to process ARP packets, the device CPU load will increase. The more frequent the packets are sent, the heavier the CPU load.

3. (Optional) Configure the interval for sending ARP detection packets.

   Run the **min-tx-interval** command to configure the interval for sending ARP detection packets.

4. Enable the ARP detection function.

   Run the **detect enable** command to enable the ARP detection function.

Step 2 Configure ARP detection for the protection port in the protection group.

Repeat Step 2(but change the working port to the protection port) to configure ARP detection for the protection port in the protection group.

Step 3 Verify ARP detection configurations at the two ports in the protection group.

Run the **display arp-detect** command to verify ARP detection configurations, such as the remote IP address and enable/disable status of ARP detection, at the two ports in the protection group.

----End

Example

Table 23-36 Data plan

Item	Value
Positions of ports in the protection group	Ports on the GIU board: 0/19/0 and 0/19/1
VLAN for ARP detection	VLAN 20
IP address of the Layer 3 interface in the VLAN	Working port: 1.1.1.2/24 Protection port: 2.2.2.2/24
Remote IP address	1.1.1.1/24 2.2.2.1/24
ARP detection times	Three times (default)
Interval for sending ARP detection packets	1s (default)

This example assumes a scenario in which the MA5600T/MA5603T/MA5608T is dual-homed to BRAS 1 and BRAS 2 through the GIU upstream board, and the "Value" column in Table 23-36 lists the data plan. When ARP detection times out, the system considers the working link interrupted and switches services to BRAS 2, ensuring uninterrupted service transmission.

To configure ARP detection in such a network scenario, do as follows:

```
//Configure the VLAN and Layer 3 interface IP address used for ARP detection of the
local device.
huawei(config)#vlan 20 smart
huawei(config)#port vlan 20 0/19 0
huawei(config)#interface vlanif 20
huawei(config-if-vlanif20)#ip address 1.1.1.2 24
huawei(config-if-vlanif20)#ip address 2.2.2.2 24 sub
huawei(config-if-vlanif20)#quit
//Configure ARP detection for the working port.
huawei(config)#arp-detect arp_test1 bind peer-ip 1.1.1.1 vlan 20 port 0/19/0
huawei(config-arp-detect-arp_test1)# detect-multiplier 3
huawei(config-arp-detect-arp_test1)# min-tx-interval 2
huawei(config-arp-detect-arp_test1)#detect enable
huawei(config-arp-detect-arp_test1)#quit
//Configure ARP detection for the protection port.
huawei(config)#arp-detect arp_test2 bind peer-ip 2.2.2.1 vlan 20 port 0/19/1
huawei(config-arp-detect-arp_test2)#detect-multiplier 3
huawei(config-arp-detect-arp_test2)#min-tx-interval 2
huawei(config-arp-detect-arp_test2)#detect enable
huawei(config-arp-detect-arp_test2)#quit
//Query configurations of the working port.
huawei(config)#display arp-detect arp_test1
 ------------------------------------------------------------------------------
  Name      : arp-test2                     Admin State : Enable
  Peerip    : 1.1.1.1                        Interval   : 2(s)
```

```
 Vlan      : 20                              Multiplier  : 3
 F/S/P     : 0/19/0                           State      : Down
 ------------------------------------------------------------------------------
//Query configurations of the protection port.
huawei(config)#display arp-detect arp_test2
 ------------------------------------------------------------------------------
 Name      : arp_test2                       Admin State : Enable
 Peerip    : 2.2.2.1                          Interval   : 2(s)
 Vlan      : 20                              Multiplier  : 3
 F/S/P     : 0/19/1                           State      : Down
 ------------------------------------------------------------------------------
```

23.20 Configuring the VoIP PSTN Service (H.248-based or MGCP-based)

This topic describes how to configure the VoIP PSTN service when the protocol adopted by the Access node is H.248 or MGCP.

Application Context

The voice over IP (VoIP) service uses the IP packet switched network for transmission after the traditional analog voice signals are compressed and packetized, to lower the cost of the voice service.

In the NGN network, the Access node functions as an access gateway (AG) and exchanges signaling with the media gateway controller (MGC) through the MG control protocol (mainly H.248 and MGCP). In this way, the VoIP, FoIP, and MoIP services are implemented under the control of the MGC. The MG interface, as an interface for the communication between the Access node (AG) and the MGC, plays a decisive role in the H.248-based or MGCP-based VoIP service. Therefore, to implement the VoIP service, the MG interface must be configured and must be in the normal state.

H.248, also called MeGaCo, is a protocol developed based on MGCP by combining the features of other media gateway control protocols. Compared with MGCP, H.248 supports more types of access technologies and supports mobility of terminals; however, the configuration of the H.248-based VoIP service is the same as that of the MGCP-based VoIP service.

Prerequisite

- According to the actual network, a route from the Access node to the MGC must be configured to ensure that the Access node and the MGC are reachable from each other.
- The voice daughter board on the control board works in the normal state.
- Electronic switch 1 must be in **location-0** (indicating that the VoIP service is supported) If the SCUB control board is used. For details about how to configure the electronic switch, see **electro-switch**.

Precaution

The MG control protocols (H.248 and MGCP) are master/slave protocols, and the MGC controls the AG to implement the call connection. Therefore, the data on the AG for

interconnection with the MGC, including the attributes of the MG interface and the attributes of the VoIP user, must be the same as the corresponding data on the MGC. Before configuring the VoIP service, you must make the data plan by considering interconnection with the MGC.

Data preparation

Table 23-37 provides the data plan for configuring the VoIP service.

Table 23-37 Data plan for configuring the H.248-based or MGCP-based VoIP service

Item			Remarks
MG interface data (The data configuration must the same as the data configuration on the MGC.)	Media and signaling parameters	Media and signaling upstream VLAN	It is used as the upstream VLAN of the VoIP service to be configured. Standard VLAN is recommended.
		Media and signaling upstream port	It is used as the upstream port of the VoIP service to be configured.
		Media and signaling IP addresses	These IP addresses must be selected from the media and signaling IP address pools. The media and signaling IP address pools consists of all the IP addresses of the Layer 3 interface of the media and signaling upstream VLAN. **NOTICE** The MGCP interface on the Access node does not support the isolation of media streams and signaling flows. Therefore, when the MGCP protocol is used, the media IP address pool and the signaling IP address pool must be the same. When the H.248 or SIP protocol is used, the media IP address pool and the signaling IP address pool can be the same or different.
		Default IP address of the MG	It is the next hop IP address from the Access node to the MGC.
	Parameters of the MG interface **NOTE** Parameters listed here are mandatory, which means that the MG interface fails to be	MG interface ID	It is the MG interface ID used for the VoIP service to be configured, which determines the virtual access gateway (VAG) specified for the user.
		Signaling port ID of the MG interface	It is the transport layer protocol port ID used for the signaling exchange between the Access node (AG) and the MGC. • Default signaling port ID defined in H.248: 2944 (text) and 2945 (binary) • Default signaling port ID defined in MGCP: 2727

Item			Remarks
	enabled if these parameters are not configured.	IP address of the primary MGC to which the MG interface belongs	When dual homing is not configured, the parameters of the primary MGC need to be configured. When dual homing is configured, the IP address and the port ID of the secondary MGC also need to be configured.
		Port ID of the primary MGC to which the MG interface belongs	
		Coding mode of the MG interface	Currently, the **text** coding mode is supported. **NOTE** For the MG interface that supports MGCP, the default coding mode is the **text** coding mode. This parameter can be queried, but cannot be configured.
		Transmission mode of the MG interface	The transmission mode is selected according to the requirements on the MGC side. Generally, UDP is adopted. **NOTE** For the MG interface that supports MGCP, the default transmission mode is UDP. This parameter can be queried, but cannot be configured.
		Domain name of the MG interface	It corresponds to the parameter **domainName** of the MG interface. • When the MGCP protocol is used, this parameter must be configured. Otherwise, the MG interface fails to be enabled. • When the H.248 protocol is used, this parameter must be configured if the parameter **MIDType** of the H.248 message is configured as the domain name. Otherwise, the MG interface fails to be enabled. In other situations, this parameter is optional.
		Profile index of the MG interface	If the MGC interconnected with the MG is also made by Huawei, set the profile index to 1 or do not set it (it is 1 by default); if the MGC interconnected with the MG is made by another manufacturer, set the profile index to the corresponding value of the manufacturer.
		Device name of the MG interface	It is supported by the H.248 protocol, and corresponds to the parameter

Item			Remarks
			deviceName of the MG interface that uses the H.248 protocol. When the H.248 protocol is used, this parameter must be configured if the parameter **MIDType** of the H.248 message is configured as the domain name. Otherwise, the MG interface fails to be enabled. In other situations, this parameter is optional.
	(Optional) Configuring the Digitmap of an MG Interface		The digitmaps here are used for certain special applications such as emergency calls and emergency standalone. The digitmaps for common phone services are issued to the AG by the MGC, and therefore such digitmaps do not need to be configured on the AG. If the services such as emergency calls and emergency standalone are not required, the digitmaps do not need to be configured.
	(Optional) Configuring the Software Parameters of an MG Interface		According to the service requirements, the configuration of software parameters determines whether the MG interface supports the functions such as dual homing and emergency standalone.
	(Optional) Configuring the Ringing Mode of an MG Interface		According to the service requirements, the ringing modes of the MG interface need to be determined.
	(Optional) Configuring the TID Format of an MG Interface		The TID format determines the generation mode of various types of user terminals on an MG interface.
Voice user data (The data configuration must be the same as the data configuration on the MGC.)	Slot of the voice service board		-
	Configuring the PSTN User Data	Phone number	The phone numbers allocated by the MGC need to be determined, and the paging numbers for users' emergency standalone need to be planned if the emergency standalone function is provided.
		TID	If the TID template with which the PSTN user is bound does not support terminal layering, this parameter needs to be configured.
		User priority	According to the service requirements, user priority needs to be specified. The user priority includes the following:

Item			Remarks
			• cat1: government1 (category 1 government users) • cat2: government2 (category 2 government users) • cat3: common (common users)
		User type	According to the service requirements, user type needs to be specified. The user type includes the following: • DEL: direct exchange lines (default) • ECPBX: earth calling PBX • LCPBX: loop calling PBX • PayPhone: pay phone
	(Optional) Configuring the System Parameters		The system parameters including the overseas version flag and message waiting indication (MWI) mode need to be configured according to the local standards to ensure that the response of the user terminal complies with the local standards.
	(Optional) Configuring the Overseas Parameters		The attributes such as the upper and lower thresholds of the flash-hooking duration need to be configured according to the local standards to ensure that the response of the user terminal complies with the local standards.
	(Optional) Configuring the Attributes of a PSTN Port		If the PSTN port needs to support the polarity reversal accounting, the PSTN port needs to be configured to support the polarity reversal pulse. Other attributes do not need to be modified if there is no special requirement.
	(Optional) Configuring the Attributes of the Ringing Current		You can adjust the ringing tone volume by modifying the attributes of the ringing current. In general, the attributes of the ringing current do not need to be modified. You need to modify the parameters of the ringing current attributes according to the local standards only when the default ringing current attributes do not meet the local standards.

Procedure

23.20.1 Configuring an MG Interface

This topic describes how to configure an MG interface for implementing the communication between the MA5600T/MA5603T/MA5608T (AG) and the MGC.

Context

- The MA5600T/MA5603T/MA5608T communicates with the MGC through either the H.248 or the MGCP protocol. One MA5600T/MA5603T/MA5608T can run only one protocol.
- One MA5600T/MA5603T/MA5608T supports up to eight MG interfaces. If a CKMC daughter board is configured in the MA5600T/MA5603T/MA5608T, up to 32 MG interfaces can be supported. Each MG interface can be configured with the interface attributes independently. Each MG interface can be configured with the attributes (such as authentication and ringing mapping) independently.

Procedure

Configuring the Upstream VLAN Interface

This topic describes how to specify the upstream VLAN interface for the media stream and the signaling flow, and how to configure the IP addresses of the Layer 3 interface. These IP addresses are the sources of the media and signaling IP address pools.

Context

Multiple IP addresses can be configured for the same VLAN Layer 3 interface. Only one IP address functions as the primary address, and other IP addresses function as the secondary addresses.

Procedure

Run the **vlan** command to add an upstream VLAN for the media stream and the signaling flow.

Step 1 Run the **port vlan** command to add the upstream ports to the VLAN.

Step 2 Configure the IP addresses of the VLAN Layer 3 interface.

1. Run the **interface vlanif** command to enter the Layer 3 interface of the upstream VLAN for the media stream and the signaling flow.
2. Run the **ip address** command to configure the IP addresses of the Layer 3 interface.

Step 3 Run the **display interface vlanif** command to check whether the IP addresses of the Layer 3 interface are the same as those in the data plan.

----End

Example

Assume that the media stream and the signaling stream are transmitted upstream through upstream port 0/19/0. To create media and signaling upstream VLAN 3 and configure IP

addresses 10.13.4.116/16 and 10.13.4.117/16 of the Layer 3 interfaces for the media IP address pool and the signaling IP address pool respectively, do as follows:

```
huawei(config)#vlan 3 standard
huawei(config)#port vlan 3 0/19 0
huawei(config)#interface vlanif 3
huawei(config-if-vlanif3)#ip address 10.13.4.116 16
huawei(config-if-vlanif3)#ip address 10.13.4.117 16 sub
huawei(config-if-vlanif3)#quit
huawei(config)#display interface vlanif 3
vlanif3 current state : UP
Line protocol current state : UP
Description : HUAWEI, SmartAX Series, vlanif3 Interface
The Maximum Transmit Unit is 1500 bytes
Internet Address is 10.13.4.116/16
Internet Address is 10.13.4.117/16 Secondary
IP Sending Frames' Format is PKTFMT_ETHNT_2, Hardware address is 00e0-fc32-1118
```

Configuring the Media and Signaling IP Address Pools

The media IP address pool defines all media IP addresses that can be used by the AG, and the signaling IP address pool defines all signaling IP addresses that can be used by the AG.

Prerequisites

The IP address of the Layer 3 interface of the media and signaling upstream VLAN must be configured. For details about how to configure the IP address, see Configuring the Upstream VLAN Interface.

Context

- The media IP address and the signaling IP address for the MG interface must be selected from the IP address pools configured here.
- The signaling IP address pool is used to store the IP addresses of the MG interfaces, and the media IP address pool is used to store the IP addresses of the media streams controlled by the signaling.
- The media IP address pool and the signaling IP address pool can be the same. Similarly, the media IP address and the signaling IP address can be the same.

NOTICE

The MGCP interface on the MA5600T/MA5603T/MA5608T does not support the isolation of media streams and signaling flows. Therefore, when the MGCP protocol is used, the media IP address pool and the signaling IP address pool must be the same. When the H.248 or SIP protocol is used, the media IP address pool and the signaling IP address pool can be the same or different.

Procedure

Run the **voip** command to enter the VoIP mode.

Step 1 Configure the media IP address pool.

1. Run the **ip address media** command to add the media IP address to the media IP address pool.

   The media IP address needs to be selected from the IP addresses of the Layer 3 interface of the media and signaling upstream VLAN.

2. Run the **display ip address media** command to check whether the media IP address pool is the same as that in the data plan.

Step 2 Configure the signaling IP address pool.

1. Run the **ip address signaling** command to add the signaling IP address to the signaling IP address pool.

   The signaling IP address needs to be selected from the IP addresses of the Layer 3 interface of the media and signaling upstream VLAN.

2. Run the **display ip address signaling** command to check whether the signaling IP address pool is the same as that in the data plan.

----End

Example

To add IP addresses 10.13.4.116/16 and 10.13.4.117/16 of Layer 3 interfaces of the media and signaling upstream VLAN to the media IP address pool and the signaling IP address pool respectively, do as follows:

```
huawei(config)#voip
huawei(config-voip)#ip address media 10.13.4.116 10.13.0.1
huawei(config-voip)#ip address media 10.13.4.117 10.13.0.1
huawei(config-voip)#ip address signaling 10.13.4.116
huawei(config-voip)#ip address signaling 10.13.4.117
huawei(config-voip)#display ip address media
  Media:
  IP Address...........: 10.13.4.116
  Subnet Mask..........: 255.255.0.0
  Gateway..............: 10.13.0.1
  MAC Address..........: 00-E0-FC-AF-91-33

  IP Address...........: 10.13.4.117
  Subnet Mask..........: 255.255.0.0
  Gateway..............: 10.13.0.1
  MAC Address..........: 00-E0-FC-AF-91-33
huawei(config-voip)#display ip address signaling
  Signaling:
  IP Address...........: 10.13.4.116
  Subnet Mask..........: 255.255.0.0
  MAC Address..........: 00-E0-FC-AF-91-33

  IP Address...........: 10.13.4.117
  Subnet Mask..........: 255.255.0.0
  MAC Address..........: 00-E0-FC-AF-91-33
```

Adding an MG Interface

This topic describes how to add an MG interface, through which the Access node can communicate with the MGC.

Context

- One Access node supports a maximum of 8 MG interfaces. If a CKMC daughter board is configured in the Access node, up to 32 MG interfaces can be supported. Each MG interface can be configured with the interface attributes independently.
- The configuration of the attributes of an MG interface is valid only to the MG interface.

Procedure

- Add an MG interface that supports H.248.
  a. Run the **display protocol support** command to query the current system protocol.
     - If the current system protocol is H.248, go to h.
     - If the current system protocol is MGCP, go to b.
  b. Run the **display if-mgcp all** command to query whether an MG interface that supports MGCP exists.
     - If such an MG interface does not exist, go to e.
     - If such an MG interface exists, go to c.
  c. Delete all configuration data of this MG interface, and then run the **shutdown(mgcp)** command to disable the MG interface.

NOTICE

This operation directly interrupts all the services on the MG interface. Hence, exercise caution when performing this operation.

  d. Run the **undo interface mgcp** command to delete the MG interface.
  e. Run the **protocol support** command to change the system protocol to H.248.
  f. Run the **save** command to save the configuration data, and then run the **reboot system** command to restart the system to make the new configuration data take effect.
  g. After the system is restarted, log in to the system, and enter the global config mode.
  h. Run the **interface h248** command to add an MG interface that supports H.248.
  i. Run the **if-h248 attribute** command to configure the attributes of the MG interface according to the data plan.
  j. Run the **display if-h248 attribute** command to check whether the attributes of the MG interface are the same as those in the data plan.
- Add an MG interface that supports MGCP.
  a. Run the **display protocol support** command to query the current system protocol.
     - If the current system protocol is MGCP, go to h.
     - If the current system protocol is H.248, go to b.
  b. Run the **display if-h248 all** command to query whether an MG interface that supports H.248 exists.

- If such an MG interface does not exist, go to e.
- If such an MG interface exists, go to c.

c. Delete all configuration data of this MG interface, and then run the **shutdown(h248)** command to disable the MG interface.

CAUTION

This operation directly interrupts all the services on the MG interface. Hence, exercise caution when performing this operation.

d. Run the **undo interface h248** command to delete the MG interface.

e. Run the **protocol support** command to change the system protocol to MGCP.

f. Run the **save** command to save the configuration data, and then run the **reboot system** command to restart the system to make the new configuration data take effect.

g. After the system is restarted, log in to the system, and enter the global config mode.

h. Run the **interface mgcp** command to add an MG interface that supports MGCP.

i. Run the **if-mgcp attribute** command to configure the attributes of the MG interface according to the data plan.

j. Run the **display if-mgcp attribute** command to check whether the attributes of the MG interface are the same as those in the data plan.

----End

Example

Assume that the MG interface ID is 0 and the H.248 protocol is used for interconnecting with the MGC. To add the MG interface, do as follows:

```
huawei(config)#display protocol support
System support H248 protocol
huawei(config)#interface h248 0
  Are you sure to add MG interface?(y/n)[n]:y
```

Assume that the MG interface ID is 0 and the MGCP protocol is used for interconnecting with the MGC. The current system protocol, however, is the H.248 protocol. It is confirmed that the H.248 interface exists in the system but is not in use. To add MG interface 0, do as follows:

```
huawei(config)#display protocol support
System support H248 protocol
huawei(config)#display if-h248 all
  ------------------------------------------------------------------------------
  MGID    TransMode State    MGPort MGIP          MGCPort MGCIP/DomainName
  ------------------------------------------------------------------------------
  0       -       Close    -     -             -      -
  ------------------------------------------------------------------------------
huawei(config)#undo interface h248 0
  Are you sure to del MG interface?(y/n)[n]:y
huawei(config)#protocol support mgcp
```

```
huawei(config)#save
huawei(config)#reboot system
  Please check whether data has saved, the unsaved data will lose if reboot
system, are you sure to reboot system? (y/n)[n]:y
```

After the system is restarted, re-log in to the system.

```
huawei(config)#display protocol support
System support MGCP protocol
huawei(config)#interface mgcp 0
  Are you sure to add MG interface?(y/n)[n]:y
```

(Optional) Configuring the Digitmap of an MG Interface

This topic describes how to configure the digitmaps for certain special applications such as emergency calls and emergency standalone. The digitmaps for common phone services are issued to the AG by the MGC, and therefore such digitmaps do not need to be configured on the AG. If the services such as emergency calls and emergency standalone are not required, the digitmaps do not need to be configured.

Prerequisites

NOTICE

The digitmap configuration is relatively complicated. The information such as the meanings and usage of the characters in a digitmap is defined in the protocol, and is not described here. This topic provides only some basic information. It is recommended that you refer to the digitmap description in ITU-T H248.1 (applicable to H.248) or RFC 3435 (applicable to MGCP) before configuring the digitmap.

Context

- A digitmap is a set of digit collection descriptors. It is a dialing scheme resident in the MG and is used for detecting and reporting digit events received on a termination. The digitmap is used to improve the efficiency of the MG in sending the callee number. That is, if the callee number matches a dialing scheme defined by the digitmap, the MG sends the callee number collectively in a message.
- A digitmap consists of strings of digits with certain meanings. When the received dialing sequence matches one of the strings, the digits are collected completely.
- To configure the emergency standalone function, you must configure the internal digitmap.

The H.248-based MG interface supports the following types of digitmaps:

- Internal digitmap
- Emergency digitmap
- Emergency call digitmap (due to call restriction in case of an overload)
- Automatic redial digitmap of the card service

Table 23-38 provides the valid characters in the strings and their meanings in the H.248 protocol. For details about the digitmap in the H.248 protocol, refer to ITU-T H248.1, which provides a better guide to the digitmap configuration.

Table 23-38 Digitmap format in the H.248 protocol

Digit or Character	Description
0-9	Indicate dialed digits 0-9.
A-D	-
E	Indicates * in the DTMF mode.
F	Indicates for # in the DTMF mode.
X	Indicates for a wildcard, indicating any digit from 0 to 9.
S	Indicates the short timer. When detecting the timer timeout, the system reports a phone number digit by digit if the phone number is detected after the dialing scheme is matched.
L	Indicates the long timer. When detecting the timer timeout, the system reports a phone number digit by digit if the phone number is detected after the dialing scheme is matched.
Z	Indicates the duration modifier, which indicates a dialing event of a long duration. It is before the event character with a fixed location. When the event duration exceeds the threshold, the dialing event fills the location.
.	Indicates that there can be multiple digits (including 0) or characters before it.
|	Used to separate the strings and indicates that each string is an optional dialing scheme.
[]	Indicates that one digit or string can be selected from the options.

The MGCP-based MG interface supports the following types of digitmaps:

- Emergency call digitmap (due to call restriction in case of an overload)
- Automatic redial digitmap of the card service

Table 23-39 provides the valid characters in the strings and their meanings in the MGCP protocol.

Table 23-39 Digitmap format in the MGCP protocol

Digit or Character	Description
0-9	Indicate dialed digits 0-9.
A-D	-
X	Indicates for a wildcard, indicating any digit from 0 to 9.

Digit or Character	Description
T	Indicates that when detecting the timer timeout, the system reports a phone number digit by digit if the phone number is detected after the dialing scheme is matched.
*	Indicates * in the DTMF mode.
#	Indicates for # in the DTMF mode.
.	Indicates that there can be multiple digits (including 0) or characters before it.
|	Used to separate the strings and indicates that each string is an optional dialing scheme.
[]	Indicates that one digit or string can be selected from the options.

Procedure

- When the system protocol is H.248, perform the following operations to configure the digitmap.
  a. In the global config mode, run the **interface h248** command to enter the H.248 mode.
  b. Run the **digitmap set** command to configure the digitmap listed in the data plan.
  c. (Optional) Run the **digitmap-timer** command to configure the digitmap timer.

     Generally, the digitmap timer is issued by the MGC. In this case, the issued digitmap timer prevails regardless of whether a timer is configured on the AG. When the MGC does not issue the digitmap timer and the default digitmap timer does not meet the service requirements, you can configure the digitmap timer in this step.
  d. Check whether the configuration of the digitmap timer is the same as that in the data plan.
     - Run the **display digitmap** command to check whether the digitmap is configured correctly.
     - Run the **display digitmap-timer** command to check whether the digitmap timer is configured correctly.
- When the system protocol is MGCP, perform the following operations to configure the digitmap.
  a. In the global config mode, run the **interface mgcp** command to enter the MGCP mode.
  b. Run the **digitmap set** command to configure the digitmap listed in the data plan.
  c. (Optional) Run the **digitmap-timer** command to configure the digitmap timer.

     Generally, the digitmap timer is issued by the MGC. In this case, the issued digitmap timer prevails regardless of whether a timer is configured on the AG. When the MGC does not issue the digitmap timer and the default digitmap timer does not meet the service requirement, you can configure the digitmap timer in this step.

d. Check whether the configuration of the digitmap timer is the same as that in the data plan.
   - Run the **display digitmap** command to check whether the digitmap is configured correctly.
   - Run the **display digitmap-timer** command to check whether the digitmap timer is configured correctly.

----End

Example

Assume that the inner digitmap of the H.248-based MG interface is configured. According to the data plan, the inner digitmap format is 1234xxxx. The digitmap timer is not configured because it is issued by the MGC. To configure the inner digitmap, do as follows:

```
huawei(config)#interface h248 0
huawei(config-if-h248-0)#digitmap set inner 1234xxxx
huawei(config-if-h248-0)#display digitmap
  ------------------------------------------------------------------------
  Inner digitmap                                  : 1234xxxx
  Emergency digitmap                              : -
  Urgent digitmap (for overload or bandwidth restrict)   : -
  Dualdial digitmap for card service                : -
  ------------------------------------------------------------------------
```

(Optional) Configuring the Software Parameters of an MG Interface

The software parameters of an MG interface mainly define certain common service attributes of the MG interface. After the configuration of the software parameters of an MG interface, the software parameters take effect immediately and the configuration is valid only to the MG interface.

Context

There are 34 software parameters (numbered from 0 to 33) of an MG interface that supports H.248. Table 23-40 lists the configurable parameters, and the other parameters are reserved in the system.

Table 23-40 Software parameters of an MG interface that supports H.248

Parameter	Description	Default Setting
2	Indicates whether the MG interface supports dual homing. To configure an MG interface to or not to support dual homing, use this parameter. If the MG interface does not support dual homing (value: 0), even if the secondary MGC is configured, the MG interface does not switch to	0: indicates that dual homing is not supported.

Parameter	Description	Default Setting
	registering with the secondary MGC when the MG interface fails to register with the primary MGC. If the MG interface supports dual homing and auto-switching (value: 2), even if the MG interface has registered with the secondary MGC, the MG interface can automatically switch back to the primary MGC when the primary MGC recovers.	
4	Indicates whether a wildcard is used for the registration of the MG interface. To configure whether a wildcard is used for the registration of an MG interface, use this parameter. Using a wildcard for registration can reduce the quantity of messages between the MG interface and the MGC. When the wildcard is not used for registration, all the terminals on the MG interface need to register with the MGC in order. The registration without a wildcard is also called "single-endpoint registration".	0: indicates that a wildcard is used.
6	Indicates whether the MG interface supports device authentication. To configure an MG interface to or not to support authentication, use this parameter. After the device authentication is supported, run the **auth(h248)** command to configure the authentication parameters, and then run the **reset(h248)** command to reset the MG interface. In this way, the MGC can manage the security of the MGs and avoid illegal registration with the MGC.	1: indicates that device authentication is not supported.

Parameter	Description	Default Setting
7	Indicates whether the MG interface supports security header. To configure an MG interface to or not to support security header, use this parameter.	1: indicates that security header is not supported.
11	Indicates whether the MG interface supports emergency standalone. To configure whether an MG interface supports emergency standalone, use this parameter. If the MG interface supports emergency standalone, the users on the MG interface can make phone calls even if the MG fails to communicate with the MGC.	0: indicates that no call is permitted.
13	Digitmap matching mode	2: indicates the minimum matching.
15	Indicates whether the function of filtering media streams by source port is enabled on an MG interface. To enable or disable the function of filtering media streams by source port on an MG interface, use this parameter. When the function of filtering media streams by source port is enabled on the MG interface, only the media streams from the specific ports can be received.	0: indicates that media streams are not filtered by source port.
16	Indicates the length of the timer for filtering the media stream source port of the MG interface. To configure the length of the timer for filtering the media stream source port of an MG interface, use this parameter. When an MG interface does not filter the source port, the MG interface automatically filters the source port if the	5s

Parameter	Description	Default Setting
	filtering timer times out.	
22	Indicates the type of the prompt tone that is played after the communication between the MG and the MGC is interrupted. To configure the type of the prompt tone after the communication between the MG and the MGC is interrupted, use this parameter.	0: indicates the busy tone.
23	Indicates the length of the timer for synchronizing the port status. To configure the length of the timer for synchronizing the port status, use this parameter.	35s
24	Indicates the maximum value of the Real-Time Transport Protocol (RTP) termination ID.	-
25	Indicates the maximum random value for the protection against avalanche of the H.248 interface.	-
26	Indicates the type of local blocking play tone.	0: indicates the busy tone.
27	Indicates the type of remote blocking play tone.	0: indicates the busy tone.
28	Indicates the duration of the howler tone.	60s
29	Indicates the duration of message waiting tone.	3s
30	Indicates the time limit of the alarm for extra long call.	60 minutes
31	Indicates whether to report the alarm for extra long call.	1: indicates that the alarm is not reported.
32	Min. auto registration interval of remotely-blocked port(s).	1800s
33	Whether MG heartbeat is shut down.	1: No, heartbeat is enabled

There are 17 software parameters (numbered from 0 to 16) of an MG interface that supports MGCP. Table 23-41 lists the configurable parameters, and the other parameters are reserved in the system.

Table 23-41 Software parameters of an MG interface that supports MGCP

Parameter	Description	Default Setting
1	Indicates whether the ongoing call is maintained if the communication between the MG interface and the MGC is abnormal. To maintain the ongoing call when the communication between the MG interface and the MGC is abnormal, use this parameter.	1: disconnects all the calls at once.
2	Indicates whether the MG interface supports dual homing. To configure whether an MG interface supports dual homing, use this parameter. If the MG interface does not support dual homing (value: 1), even if the secondary MGC is configured, the MG interface does not switch to register with the secondary MGC when the MG interface fails to register with the primary MGC.	0: indicates that dual homing is supported.
3	Indicates whether the heartbeat message between the MG and the MGC is disabled. To configure whether the heartbeat message between the MG and the MGC is disabled, use this parameter.	1: indicates that the heartbeat message is not disabled.
4	Indicates whether a wildcard is used for the registration of the MG interface. To configure whether a wildcard is used for the registration of an MG interface, use this parameter. Using a wildcard for registration can reduce the quantity of messages between the MG interface and the	0: indicates that a wildcard is used.

Parameter	Description	Default Setting
	MGC. When a wildcard is not used for registration, all the terminals on the MG interface need to register with the MGC in order. The registration without a wildcard is also called "single-endpoint registration".	
5	Indicates the MGC type. To select the MGC of a different type, use this parameter.	0
6	Indicates the maximum time threshold for responding to the heartbeat messages. To configure the maximum times for transmitting the heartbeat message continuously, use this parameter.	3
7	Indicates whether to report the heartbeat with the MG as an endpoint. To configure whether to report the heartbeat with the MG as an endpoint, use this parameter.	0: indicates that reporting the heartbeat with the MG as an endpoint is not supported.
10	Indicates the point-to-point (P2P) fault reporting. To configure whether the MG interface supports the P2P fault reporting from the ISDN terminals, use this parameter.	0: indicates that the P2P fault is reported.
11	Indicates the point-to-multipoint (P2MP) fault reporting. To configure whether the MG interface supports the P2MP fault reporting from the ISDN terminals, use this parameter.	1: indicates that the P2MP fault is not reported.
12	Indicates the type of local blocking play tone.	0: indicates the busy tone.
13	Indicates the type of remote blocking play tone.	0: indicates the busy tone.
14	Indicates the RTP filtering	1: indicates that the RTP

Parameter	Description	Default Setting
	switch of the MGCP interface. To configure whether the RTP filtering function is enabled, use this parameter. When the RTP filtering function is enabled, only the media stream from the specific ports can be received.	filtering function is not enabled.
15	Indicates the duration of the howler tone.	60s
16	Whether the timer symbol "T" follows the number string reported by the signaling.	0: Yes

Procedure

- When the system protocol is H.248, perform the following operations to configure the software parameters of an MG interface.
  a. In the global config mode, run the **interface h248** command to enter the MG interface mode.
  b. Run the **mg-software parameter** command to configure the software parameters listed in the data plan.
  c. Run the **display mg-software parameter** command to check whether the software parameters of the MG interface are the same as those in the data plan.
- When the system protocol is MGCP, perform the following operations to configure the software parameters of an MG interface.
  a. In the global config mode, run the **interface mgcp** command to enter the MG interface mode.
  b. Run the **mg-software parameter** command to configure the software parameters listed in the data plan.
  c. Run the **display mg-software parameter** command to check whether the software parameters of the MG interface are the same as those in the data plan.

----End

Example

To configure software parameter 11 of H.248-based MG interface 0 to 1 so that the MG interface supports emergency standalone and allows only internal calls, do as follows:

```
huawei(config)#interface h248 0
huawei(config-if-h248-0)#mg-software parameter 11 1
huawei(config-if-h248-0)#display mg-software parameter 11
  ------------------------------------------------
  Interface Id:0         para index:11  value:1
  ------------------------------------------------
 APPENDIX:
```

```
-----------------------------------------------
 Interface software parameter name:
 11: Stand alone support
    0: None
    1: Inner
    2: Emergency
    3: Both
```

(Optional) Configuring the Ringing Mode of an MG Interface

This topic describes how to configure the ringing mode of an MG interface to meet different user requirements.

Procedure

- If the system protocol is H.248, perform the following operations to configure the ringing mode of an MG interface.
  a. Check whether the preset ringing mode in the system meets the requirements according to the Usage Guidelines of the **mg-ringmode add** command.
     - If the preset ringing mode meets the requirements, go to c.
     - If the preset ringing mode does not meet the requirements, proceed to b.
  b. In the global config mode, run the **user defined-ring modify** command to configure the break-make ratio of user-defined ringing mode to form a ringing mode that meets the user requirement.

NOTICE

- After configuring the user-defined ringing mode, you need to reset the corresponding service board to make the configuration data take effect. Thus, the user of the service board can use the new user-defined ringing mode. If the service board has been in service for a period, evaluate the impact on the online users before resetting the service board, and then determine whether to perform this operation.
- The system supports 16 user-defined ringing modes, which can be modified but cannot be added or deleted.

  c. Run the **interface h248** command to enter the H.248 mode.
  d. Run the **mg-ringmode add** command to add a ringing mapping.

NOTICE

1. The parameter *mgcpara* on the MG must be the same as the parameter *mgcpara* on the MGC.
2. User-defined ringing modes 0 to 15 correspond to cadence ringing modes 128 to 143 respectively, and correspond to initial ringing modes 144 to 159 respectively. For example, if the cadence ringing mode is 128, user-defined ringing mode 0 is selected. If the initial ringing mode is 144, user-defined ringing mode 0 is selected.

e. Run the **display mg-ringmode attribute** command to check whether the ringing mapping is the same as that in the data plan.

- If the system protocol is MGCP, perform the following operations to configure the ringing mode of an MG interface.
  a. According to the Usage Guidelines of the **mg-ringmode add** command, check whether the preset ringing mode in the system meets the requirement.
     - If the preset ringing mode meets the requirement, go to c.
     - If the preset ringing mode does not meet the requirement, proceed to b.
  b. In the global config mode, run the **user defined-ring modify** command to configure the user-defined ringing mode.

NOTICE

After configuring the user-defined ringing mode, you need to reset the corresponding service board to make the configuration data take effect. Thus, the user of the service board can use the new user-defined ringing mode. If the service board has been in service for a period, evaluate the impact on the online users before resetting the service board, and then determine whether to perform this operation.

c. Run the **interface mgcp** command to enter the MGCP mode.

d. Run the **mg-ringmode add** command to add a ringing mapping.

NOTICE

The parameter *mgcpara* on the MG must be the same as the parameter *mgcpara* configured on the MGC.

e. Run the **display mg-ringmode attribute** command to check whether the ringing mapping is the same as that in the data plan.

----End

Example

Assume that the MG interface ID is 0, the peer parameter ID issued by the MGC is 0, the cadence ringing is 1:4 (the value of the corresponding parameter *cadence* is 0), and the initial

ringing is 1:2 (the value of the corresponding parameter *initialring* is 17). To configure the ringing mode of MG interface 0, do as follows:

```
huawei(config)#interface h248 0
huawei(config-if-h248-0)#mg-ringmode add 0 0 17
huawei(config-if-h248-0)#display mg-ringmode attribute
{ <cr>|mgcpara<U><0,15> }:

  Command:
         display mg-ringmode attribute
  ------------------------------------------------------
   MGID      PeerPara   CadenceRing   InitialRing
  ------------------------------------------------------
   0         0          0             17
  ------------------------------------------------------
```

Assume that the MG interface ID is 0, the peer parameter ID issued by the MGC is 0, the break-make ratio of user-defined ringing mode 0 is 0.4sec On, 0.2sec Off, 0.4sec On, 2.0sec Off, and the initial ringing and the cadence ringing use user-defined ringing mode 0 (the values of the corresponding parameters *cadence* and *initialring* are 128 and 144 respectively). To configure the ringing mode of MG interface 0, do as follows:

```
huawei(config)#user defined-ring modify 0 para1 400 para2 200 para3 400 para4 2000
  Note: Please reset the service board to make configured parameter be valid
huawei(config)#display user defined-ring
  -------------------------------------------------
  RingType Para1 Para2 Para3 Para4 Para5 Para6
  -------------------------------------------------
  0        400   200   400   2000    0     0
  1          0     0     0     0     0     0
  2          0     0     0     0     0     0
  3          0     0     0     0     0     0
......
  14         0     0     0     0     0     0
  15         0     0     0     0     0     0
  -------------------------------------------------
huawei(config)#interface h248 0
huawei(config-if-h248-0)#mg-ringmode add 1 128 144
huawei(config-if-h248-0)#display mg-ringmode attribute
{ <cr>|mgcpara<U><0,15> }:
{ <cr>|mgcpara<U><0,15> }:

  Command:
         display mg-ringmode attribute
  ------------------------------------------------------
   MGID      PeerPara   CadenceRing   InitialRing
  ------------------------------------------------------
   0         1          128           144
  ------------------------------------------------------
```

(Optional) Configuring the TID Format of an MG Interface

The TID format determines how various user terminal IDs on the MG interface are generated.

Prerequisites

NOTICE

The configuration of the TID format is relatively complicated. The information such as the syntax rules with which the terminal prefix must comply and the requirements for the character string of the TID template is defined in the protocol, and is not described here. This topic describes only some basic information. It is recommended that you refer to the TID description in ITU-T H248.1 (applicable to H.248) or RFC 3435 (applicable to MGCP) before configuring the TID format.

Context

The TID format consists of the terminal prefix and the TID template. The TID template defines the generation mode of the TID excluding the terminal prefix. A TID consists of a terminal prefix and a character string generated by a TID template.

The TID templates that are bound to various types of users on the MG interface determine whether the users support terminal layering.

- If the parameter list of the TID template includes only keyword "G", the TID template is used by the non-layering users. Users bound with this template do not support terminal layering.
- If the parameter list of the TID template includes only keywords "F", "S", "P", "B" ("B" is not available to PSTN users), the TID template is used by the layering users. Users bound with this template support terminal layering.

NOTE

The meaning of each keyword is as follows:

- F indicates the subrack ID.
- S indicates the slot ID.
- P indicates the port ID.
- B indicates the B channel ID (only for ISDN BRA and ISDN PRA terminals).
- G indicates the general permanent termination index.
- R indicates the RTP ephemeral termination index (only for the RTP ephemeral termination, which exists only when the system protocol is H.248. This termination is not involved unless special remarks are provided.)

When adding a user that supports terminal layering, you cannot and do not have to specify the parameter **terminalid** because the system automatically allocates a terminal ID. When adding a user that does not support terminal layering, you must specify the parameter **terminalid**.

You can run the **display tid-format(h248)** command or **display tid-format(mgcp)** command to query the TID formats of various types of users on the current MG interface. In the query result, **template-index** indicates the index of the TID template that is bound to the type of users. Then, run the **display tid-template** command to check whether the TID template supports the layering configuration. Hence, you can check whether the user supports terminal layering.

Precaution

- There are 19 default TID templates (templates 0-18) in the system. The default TID templates can be referenced, but cannot be added, modified, or deleted.
- The configuration of terminal layering on the MG must be the same as that on the MGC.
- If certain type of terminals exists on an interface and the interface is not disabled, the terminal prefix of this type of terminals cannot be modified.
- If a certain type of terminals exists on an interface, the TID format (including the terminal prefix and the index of the TID template) of this type of terminals cannot be changed.
- The terminal prefix must comply with the following syntax rules: The prefix must not exceed 64 characters. Only letters, digits, "_", and "/" are the characters allowed for input. The first character must be a letter, and the last character must not be a digit.
- The length of the TID, which is generated by combining the TID template and the terminal prefix that you configured, must not exceed 64 characters.

Procedure

- If the system protocol is H.248, to configure the TID format on the current MG interface, do as follows:
  a. Run the **display tid-template** command to query the information about the default TID template of the system.
  b. If the default TID template cannot meet the service requirements, run the **tid-template add** command to add a TID template that meets the service requirements. If the default TID template meets the service requirements, proceed to c.
  c. Run the **interface h248** command to enter the H.248 mode.
  d. Configure the TID templates and the terminal prefix of various types of users on the current MG (VAG).

NOTICE

The terminal prefix of PSTN, BRA and PRA must be all identical or unique.

    - In the H.248 mode, run the **tid-format rtp** command to configure the TID template and the terminal prefix of the RTP ephemeral termination.
    - In the H.248 mode, run the **tid-format pstn** command to configure the TID template and the terminal prefix of the PSTN user.
    - In the H.248 mode, run the **tid-format bra** command to configure the TID template and the terminal prefix of the ISDN BRA user.
    - In the H.248 mode, run the **tid-format pra** command to configure the TID template and the terminal prefix of the ISDN PRA user.

  e. Run the **display tid-format(h248)** command to check whether the TID templates and the terminal prefixes of various types of users on the current MG interface are the same as those in the data plan.
- If the system protocol is MGCP, to configure the TID format on the current MG interface, do as follows:

a. Run the **display tid-template** command to query the information about the default TID template of the system.
b. If the default TID template cannot meet the service requirements, run the **tid-template add** command to add a TID template that meets the service requirements. If the default TID template meets the service requirements, proceed to c.
c. Run the **interface mgcp** command to enter the MGCP mode.
d. Configure the TID templates and the terminal prefix of various types of users on the current MG (VAG).
   - In the MGCP mode, run the **tid-format pstn** command to configure the TID template and the terminal prefix of the PSTN user.
   - In the MGCP mode, run the **tid-format bra** command to configure the TID template and the terminal prefix of the ISDN BRA user.
   - In the MGCP mode, run the **tid-format pra** command to configure the TID template and the terminal prefix of the ISDN PRA user.
e. Run the **display tid-format(mgcp)** command to check whether the TID templates and the terminal prefixes of various types of users on the current MG interface are the same as those in the data plan.

----End

Example

Assume that in the H.248 mode, the terminal prefix of the PSTN user on MG interface 1 is aln/, and the layering TID template 3 is used. To add a PSTN user on port 0/2/0 and check whether the system automatically allocates a TID generated according to the template, do as follows:

```
huawei(config)#display tid-template 3//Query the information about TID template 3
  -------------------------------------------------
  Index     : 3
  Format    : %u/%u/%u
  Para-list  : F+1,S+1,P+1  //The parameter list of the TID template includes keyword
"F", "S",
                //and "P", which indicates that this template supports terminal layering.
  Name      : Aln_Not_Fixed_1
  -------------------------------------------------

huawei(config)#interface h248 1
huawei(config-if-h248-1)#tid-format pstn prefix aln template 3
huawei(config-if-h248-1)#quit
huawei(config)#esl user
huawei(config-esl-user)#mgpstnuser add 0/2/0 1
huawei(config-esl-user)#display mgpstnuser 0/2/0
{ <cr>|endframeid/slotid/portid<S><1,15> }:

  Command:
         display mgpstnuser 0/15/0
  ------------------------------------------------------------------------
  F  /S /P    MGID    TelNo          Priority PotsLineType TerminalID
  ------------------------------------------------------------------------
  0  /2 /0    1       -              Cat3     DEL         aln/1/3/1
           //The system allocates the terminal ID according to the TID format.
```

```
------------------------------------------------------------------------
```

Enabling an MG Interface

Enabling an MG interface is to reset an MG interface to make the MG interface register with the MGC (or to make the modified attributes of the MG interface take effect) after the configuration of the MG interface is complete, so that the MG interface can work in the normal state.

Precaution

CAUTION

For the MG interface that has been in service, this operation interrupts the ongoing services carried on the MG interface. Hence, exercise caution when performing this operation.

Procedure

- Enable the MG interface that adopts the H.248 protocol.
  a. Run the **interface h248** command to enter the H.248 mode.
  b. Run the **reset coldstart** command to enable the MG interface.
  c. Run the **quit** command to return to the global config mode, and then run the **display if-h248 all** command to check whether the MG interface is in the normal state.
- Enable the MG interface that adopts the MGCP protocol.
  a. Run the **interface mgcp** command to enter the MGCP mode.
  b. Run the **reset** command to enable the MG interface.
  c. Run the **quit** command to return to the global config mode, and then run the **display if-mgcp all** command to check whether the MG interface is in the normal state.

----End

Example

To enable H.248-based MG interface 0, do as follows:

```
huawei(config)#interface h248 0
huawei(config-if-h248-0)#reset coldstart
  Are you sure to reset MG interface?(y/n)[n]:y
huawei(config-if-h248-0)#quit
huawei(config)#display if-h248 all
------------------------------------------------------------------------------
MGID    TransMode State        MGPort MGIP         MGCPort MGCIP/DomainName
------------------------------------------------------------------------------
0       UDP      Normal       2944   10.10.10.11    2944   10.10.20.11
------------------------------------------------------------------------------
```

To enable MGCP-based MG interface 0, do as follows:

```
huawei(config)#interface mgcp 0
huawei(config-if-mgcp-0)#reset
  Are you sure to reset MG interface?(y/n)[n]:y
huawei(config-if-mgcp-0)#quit
huawei(config)#display if-mgcp all
  ------------------------------------------------------------------------------
  MGID    State         MGPort MGIP           MGCPort MGCIP/DomainName
  ------------------------------------------------------------------------------
  0       Normal        2727   10.10.10.11    2727    10.10.20.11
  1       Wait ack      2527   10.10.10.12    2727    10.10.20.12
  ------------------------------------------------------------------------------
```

23.20.2 Configuring the VoIP PSTN User

After an MG interface is configured, you can add plain old telephone service (POTS) users on the MG interface to implement the VoIP PSTN service.

Procedure

Configuring the PSTN User Data

This topic describes how to configure the PSTN user data (the same as the corresponding data on the MGC) on the MG interface to provide the POTS terminal with the access to the network for VoIP service.

Prerequisites

The POTS service board must be inserted into the planned slot, and the RUN ALM LED of the board must be green and must be on for 1s and off for 1s repeatedly.

NOTE

You can add a service board in two ways (see the Usage Guideline of the **board add** command). It is recommended that you insert the service board into the planned slot and then confirm the board.

Procedure

In the global config mode, run the **board confirm** command to confirm the service board.

Step 1 Add a PSTN user.

1. In the global config mode, run the **esl user** command to enter the ESL user mode.
2. Run the **mgpstnuser add** or **mgpstnuser batadd** command to add a PSTN user or add PSTN users in batches.

NOTICE

- When you add a PSTN user, the terminal ID must be configured and must be different from the terminal ID of an existing PTSN user if the TID template with which the PSTN user on the MG interface is bound is not a layering template.
- When you add a PSTN user, the configuration of the terminal ID is not required and the system automatically allocates the terminal ID if the TID template with which the PSTN user on the MG interface is bound is a layering template.
- When adding a PSTN user, you can configure the phone number (parameter *telno*). The phone number configured, however, can be used only as the paging number for emergency standalone. Phone numbers for normal call services are allocated by the MGC. It is recommended that the phone number configured here be the same as the phone number allocated by the MGC. In addition, the phone number must be unique in the MG. This is to avoid the number conflict that may occur when emergency standalone is enabled. If this parameter is not set, the phone number is null by default.
- For details about the relation between the TID template and the terminal layering, see the Background Information in (Optional) Configuring the TID Format of an MG Interface.

3. Run the **display mgpstnuser** command to check whether the PSTN user data is the same as that in the data plan.

Step 2 (Optional) Configure the attributes of the PSTN user.

The attributes of a PSTN user need to be configured when the default settings are not consistent with the actual application.

1. Run the **mgpstnuser attribute set** or **mgpstnuser attribute batset** command to configure the attributes of the PSTN user.
2. Run the **display mgpstnuser attribute** command to check whether the attributes of the PSTN user are the same as those in the data plan.

----End

Example

Assume that the phone numbers of 32 PSTN users are 83110000-83110031, the **terminalid** values are 0-31 (the TID template to which the PSTN users under the MG interface are bound does not support layering and **terminalid** should be allocated manually), and the default values are used for other attributes. To add the 32 PSTN users in slot 0/2 under MG 0 in batches, do as follows:

```
huawei(config)#board confirm 0/2
huawei(config)#esl user
huawei(config-esl-user)#mgpstnuser batadd 0/2/0 0/2/31 0 terminalid 0 telno 83110000
huawei(config-esl-user)#display mgpstnuser 0 0 32
  ---------------------------------------------------------------------------
  F  /S /P    MGID    TelNo          Priority PotsLineType TerminalID
  ---------------------------------------------------------------------------
  0  /2 /0    0       83110000       Cat3    DEL        A0
  0  /2 /1    0       83110001       Cat3    DEL        A1
  ......
  0  /2 /30   0       83110030       Cat3    DEL        A30
  0  /2 /31   0       83110031       Cat3    DEL        A31
  ---------------------------------------------------------------------------
```

(Optional) Configuring the System Parameters

This topic describes how to configure the system parameters including the overseas version flag and message waiting indication (MWI) mode according to the local standards to ensure that the response of the user terminal complies with the local standards.

Procedure

Run the **system parameters** command to configure the system parameters.

Step 1 Run the **display system parameters** command to check whether the system parameters are the same as those in the data plan.

----End

Example

To configure the overseas version flag (system parameter 1) to Hong Kong (parameter value 1), do as follows:

```
huawei(config)#system parameters 1 1
huawei(config)#display system parameters 1
  ---------------------------------------------------------------------------
  Parameter name index: 1    Parameter value: 1
  Mean: Overseas version flag, 0:China, 1:HongKong, 2:Brazil, 3:Egypt, 4:
Singapore, 5:Thailand, 6:France, 7:Britain MSFUK, 8:Britain ETSI, 9:Bulgaria,
10:Reserved, 11:Austria, 12:Hungary, 13:Poland
  ---------------------------------------------------------------------------
```

(Optional) Configuring the Overseas Parameters

By default, the overseas parameters are configured according to the Chinese standards. In the actual service configuration, the attributes such as the upper and lower thresholds of the flash-hooking duration can be configured according to the local standards to ensure that the response of the user terminal complies with the local standards.

Procedure

Run the **oversea parameters** command to configure the overseas parameters.

Step 1 Run the **display oversea parameters** command to check whether the overseas parameters are the same as those in the data plan.

----End

Example

To set the upper flash-hooking threshold (overseas feature parameter 0) to 800 ms (in compliance with the Hong Kong standard) and the lower flash-hooking threshold (overseas feature parameter 1) to 100 ms (in compliance with the Hong Kong standard), do as follows:

```
huawei(config)#oversea parameters 0 800
huawei(config)#oversea parameters 1 100
huawei(config)#display oversea parameters
```

```
{ <cr>|name<U><0,99> }:

 Command:
       display oversea parameters
 ------------------------------------------------------------------------------
 Parameter name index: 0    Parameter value: 800
 Mean: Hooking upper threshold(ms), reference: China:350, HongKong:800
 ------------------------------------------------------------------------------
 Parameter name index: 1    Parameter value: 100
 Mean: Hooking lower threshold(ms), reference: China:100, HongKong:100
 ------------------------------------------------------------------------------
 Parameter name index: 2    Parameter value: 0
 Mean: Flag of applying PARKED LINE FEED or not when user port is locked, 0:not
apply, 1:apply
 ------------------------------------------------------------------------------
 Parameter name index: 3    Parameter value: 0
 Mean: The detect time of flash upper limit to onhook, default value: 0ms
 ------------------------------------------------------------------------------
 Parameter name index: 4    Parameter value: 0
 Mean: DTMF Detector Tuning, 0:Q.24, 1:Russia
------------------------------------------------------------------------------
 Parameter name index: 5    Parameter value: 0
 Mean: Flag of changing TEI value for software switch, 0:no change, 1:set TEI
value to 1, 2:set TEI value to 0. Default: 0
 ------------------------------------------------------------------------------
```

(Optional) Configuring the Attributes of a PSTN Port

This topic describes how to configure the attributes of a PSTN port to ensure that the PSTN port can meet the actual application requirements.

Context

The MA5600T/MA5603T/MA5608T supports the following attributes of a PSTN port:

- Physical attributes (including whether to support the polarity reversal pulse, whether to support the port locking, and dialing mode). For details about how to configure the physical attributes of a PSTN port, see **pstnport attribute set**.
- Electrical attributes (including the impedance and the current). For details about how to configure the electrical attributes of a PSTN port, see **pstnport electric set**.
- KC attributes (including the KC accounting mode and the valid voltage). For details about how to configure the KC attributes of a PSTN port, see **pstnport kc set**.

Procedure

In the global config mode, run the **pstnport** command to enter the PSTN port mode.

Step 1 Run the **pstnport attribute batset** or **pstnport attribute set** command to configure the physical attributes of the PSTN port.

Step 2 Run the **pstnport electric batset** or **pstnport electric set** command to configure the electrical attributes of the PSTN port.

Step 3 Run the **pstnport kc batset** or **pstnport kc set** command to configure the KC attributes of the PSTN port.

Step 4 Check whether the attribute configuration of the PSTN port is the same as that in the data plan.

- Run the **display pstnport attribute** command to query the physical attributes of the PSTN port.
- Run the **display pstnport electric** command to query the electrical attributes of the PSTN port.
- Run the **display pstnport kc** command to query the KC attributes of the PSTN port.

----End

Example

To configure the 32 PSTN ports of the board in slot 0/3 to support the polarity reversal accounting, do as follows:

```
huawei(config)#pstnport
huawei(config-pstnport)#pstnport attribute batset 0/3/0 0/3/31 reverse-pole-pulse
enable
huawei(config-pstnport)#display pstnport attribute 0/3
  ------------------------------------------------------------------------
  F  /S  /P            0/3 /0
  ReversePolepulse       Enable
  PulseLevel            100(ms)
  PolarityReverseMode    Hard-polarity-reverse
  Dial-Mode             DTMF-Pulse-Both
  LineLock              Enable
  NlpMode               Nlp normal mode
  PolarityReverseWhenCLIP Disable
  PulsePeriodUpperLimit   200(ms)
  PulsePeriodLowerLimit   50(ms)
  PulseDurationUpperLimit 90(ms)
  PulseDurationLowerLimit 30(ms)
  PulsePauseUpperLimit    90(ms)
  PulsePauseLowerLimit    30(ms)
  OffhookTime(Idle)      80(ms)
  OffhookTime(Ring)      200(ms)
  OffhookTime(Fsk)       50(ms)
  ------------------------------------------------------------------------
  F  /S  /P            0/3 /1
  ReversePolepulse       Enable
  PulseLevel            100(ms)
  PolarityReverseMode    Hard-polarity-reverse
  Dial-Mode             DTMF-Pulse-Both
  LineLock              Enable
  NlpMode               Nlp normal mode
  PolarityReverseWhenCLIP Disable
  PulsePeriodUpperLimit   200(ms)
  PulsePeriodLowerLimit   50(ms)
  PulseDurationUpperLimit 90(ms)
  PulseDurationLowerLimit 30(ms)
  PulsePauseUpperLimit    90(ms)
```

```
PulsePauseLowerLimit   30(ms)
OffhookTime(Idle)      80(ms)
OffhookTime(Ring)      200(ms)
OffhookTime(Fsk)       50(ms)
------------------------------------------------------------------------------
F  /S  /P            0/3 /31
ReversePolepulse       Enable
PulseLevel            100(ms)
PolarityReverseMode    Hard-polarity-reverse
Dial-Mode             DTMF-Pulse-Both
LineLock              Enable
NlpMode               Nlp normal mode
PolarityReverseWhenCLIP Disable
PulsePeriodUpperLimit   200(ms)
PulsePeriodLowerLimit   50(ms)
PulseDurationUpperLimit 90(ms)
PulseDurationLowerLimit 30(ms)
PulsePauseUpperLimit   90(ms)
PulsePauseLowerLimit   30(ms)
OffhookTime(Idle)      80(ms)
OffhookTime(Ring)      200(ms)
OffhookTime(Fsk)       50(ms)
------------------------------------------------------------------------------
```

NOTE

When a call begins and ends, the MG shows the start time and the end time based on the polarity reversal on the subscriber line. The billing terminal that supports the polarity reversal accounting function, such as a charging phone set, implements the polarity reversal accounting function based on the start time and the end time of a call.

(Optional) Configuring the Attributes of the Ringing Current

You can adjust the ringing volume and ringing tone by modifying the attributes of the ringing current. In general, the attributes of the ringing current do not need to be modified. You need to modify the attributes of the ringing current according to the local standards only when the default ringing current attributes do not meet the local standards.

Context

The attributes of the ringing current include the following two parameters:

- Ringing current frequency: A higher frequency indicates a sharper ringing tone. The default ringing current frequency is 25 Hz.
- AC amplitude (AC voltage): A greater amplitude indicates a louder ringing tone. The default AC amplitude is 75 Vrms.

Procedure

In the global config mode, run the **voip** command to enter the VoIP mode.

Step 1 Run the **ring attribute set** command to configure the attributes of the ringing current according to the data plan.

Step 2 Run the **display ring attribute** command to check whether the attributes of the ringing current are the same as those in the data plan.

----End

Example

To set the ringing current frequency to 50 Hz (parameter value 2), AC amplitude to 50 Vrms (parameter value 2), do as follows:

```
huawei(config-voip)#ring attribute set frequency 2 acamplitude 2
huawei(config-voip)#display ring attribute
  ringing current frequency  : 50HZ
  ringing current acamplitute: 50VRMS
```

23.21 Configuring the VoIP ISDN BRA Service (H.248-based)

This topic describes how to configure the VoIP ISDN BRA service on an IP network. When the Access node uses the H.248 protocol, the device supports the access of the ISDN BRA user. ISDN technology provides end-to-end (E2E) digital connection and supports multiple types of voice and non-voice telecom services.

Prerequisites

According to the actual network, a route from the Access node to the MGC must be configured to ensure that the Access node communicates with the MGC normally.

Context

- The ISDN is integrated services digital network. Defined by the International Telegraph and Telephone Consultative Committee (CCITT), the ISDN technology provides two types of user-network interfaces based on the user-network interface reference model.
  - ISDN basic rate interface (BRI), which supports a rate of 144 kbit/s and provides two B channels (for carrying services) and one D channel (for transmitting call control signaling and maintenance and management signaling). The rate of B channels is 64 kbit/s and the rate of D channel is 16 kbit/s. The service carried on the ISDN BRI is the ISDN basic rate access (BRA) service.
  - ISDN primary rate interface (PRI), which supports a rate of 2.048 Mbit/s and provides 30 B channels and one D channel. The rates of the B channel and D channel are both 64 kbit/s. The service carried on the ISDN PRI is the ISDN primary rate access (PRA) service.

Precaution

The media gateway control protocol (MGCP) is a master/slave protocol. under which the MGC controls the AG to implement call connection and disconnection. The data on the AG for the interconnection with the MGC, such as the attributes of the MG interface and the attributes of the VoIP user, must be the same as the corresponding data on the MGC. Therefore, before configuring the VoIP service, you must contact MGC engineers to check and ensure that the interconnection data plan for the AG is consistent with the corresponding plan for the MGC.

Data preparation

Table 23-42 provides the data plan for configuring the H.248-based VoIP ISDN BRA service.

Table 23-42 Data plan for configuring the H.248-based VoIP ISDN BRA service

Item			Remarks
MG interface data (The data must be consistent with the data on the MGC.)	Parameters of the media stream and signaling stream	Upstream VLAN for media and signaling streams	It is used as the upstream VLAN of the VoIP service to be configured. Standard VLAN is recommended.
		Uplink port for media and signaling streams	It is used as the uplink port for the VoIP service to be configured.
		Media IP address and signaling IP address	These IP addresses must be contained in the media and signaling IP address pools. The media and signaling IP address pools consist of all the IP addresses of the L3 interface of the upstream VLAN for media and signaling streams.
		Default IP address of the MG	It is the next hop IP address from the Access node to the MGC.
	Attribute parameters of the MG interface **NOTE** There are many MG interface parameters. Only mandatory parameters are listed here. If the mandatory parameters are not configured, the MG interface cannot be started.	MG interface ID	It is the ID of the MG interface used by the VoIP service to be configured.
		Signaling port ID of the MG interface	It is the transport layer protocol port ID used for the signaling exchange between the Access

Item			Remarks
			node (AG) and the MGC. The default signaling port ID defined in H.248 is 2944 (text) and 2945 (binary).
		IP address of the MGC to which the MG interface belongs	The MGC can be specified by IP address or the domain name. The IP address is adopted here. When dual homing is not configured, you can configure the parameters of only the primary MGC. When dual homing is configured, you also need to configure the IP address and port ID of the secondary MGC.
		Port ID of the MGC to which the MG interface belongs	
		Domain name of the MGC to which the MG interface belongs	
		Coding mode of the MG interface	Only the **text** mode is supported.
		Transmission mode of the MG interface	The transmission mode is selected according to the requirements of the MGC. Generally, UDP is used.
		Domain name of the MG interface	It corresponds to the **domainName** parameter of the MG interface. When the H.248 protocol is used, this parameter must be configured if the **MIDType** parameter of the H.248 message is configured to **domainName**. Otherwise, the MG interface cannot be started. In other situations, this parameter is optional.
		Device name of the MG interface	It is supported by only the H.248 protocol, and it corresponds to the **deviceName** parameter of the MG interface that uses the H.248 protocol. This parameter must be configured if the **MIDType** parameter of the H.248 message is configured to **domainName**.

Item			Remarks
			Otherwise, the MG interface cannot be started. In other situations, this parameter is optional.
	(Optional) Configuring the Software Parameters of an MG Interface		Whether the MG interface supports the functions such as dual homing and emergency standalone is determined by the service requirements.
	(Optional) Configuring the TID Format of an MG Interface		The TID format determines the generation mode of various types of user terminals on an MG interface.
IUA link	Adding an IUA Link Set		The IUA link can be configured only after the IUA link set is configured.
	Adding an IUA Link **NOTE** The local port ID, local IP address, remote port ID, and remote IP address of different links must not be completely same; otherwise, the service cannot be configured.	IUA link ID	It indicates the link for transmitting the signaling.
		IUA link set ID	-
		Local port ID	To activate the link normally, it must be the same as the remote port ID configured on the MGC.
		Local IP address	It must be the same as the remote IP address of the link configured on the MGC. In addition, the local IP addresses of the links that are in the same link set must be the same. (The IP address must exist in the media IP address pool.)
		Remote port ID	To activate the link normally, it must be the same as the local port ID configured on the MGC.
		Remote IP address	It must be the same as the local IP address of the link configured on the MGC. The SCTP protocol supports the multi-homing function. That is, one link can be configured with the IP addresses of multiple MGCs as the remote IP addresses. When one MGC is faulty, the link can be switched to other MGCs automatically. This ensures that the service is not affected. The Access node

Item			Remarks
			supports the configuration of the active remote IP address and standby remote IP address.
ISDN BRA user data (The data must be consistent with the data on the MGC.)	Slot that houses the ISDN BRA service board		-
	Configuring the ISDN BRA User Data	Termination ID	If the TID format bound to the BRA user does not support terminal layering function, this parameter needs to be configured, and the configuration must be consistent with the configuration on the MGC.
		User priority	The user priority must be specified according to the service requirements. There are three categories of user priorities, which are as follows: • cat1: government1 (category 1 government user) • cat2: government2 (category 2 government user) • cat3: normal (common user, default) The priorities of cat1, cat2, and cat3 are in descending order. Without special requirements, the default cat3 is adopted.
		Interface ID	It indicates the interface for the BRA user data to pass through the MG and MGC. The configuration of this parameter must be consistent with the corresponding configuration on the MGC.
	(Optional) Configuring the System Parameters		The system parameters including the oversea version flag and message waiting indication (MWI) mode need to be configured according to the local standards to ensure that the response of the user terminal complies with the local standards.
	(Optional) Configuring the Overseas Parameters		The attributes such as the upper and lower thresholds of the flash-hooking duration need to be configured according to the

Item		Remarks
		local standards to ensure that the response of the user terminal complies with the local standards.
	(Optional) Configuring the Attributes of an ISDN BRA Port	The attributes such as the working mode, remote power supply status, and auto-deactivation status of the port can be configured. Modify such attributes only if there is a special requirement.

23.21.1 Configuring an MG Interface

This topic describes how to configure an MG interface for implementing the communication between the MA5600T/MA5603T/MA5608T (AG) and the MGC.

Context

- The MA5600T/MA5603T/MA5608T communicates with the MGC through either the H.248 or the MGCP protocol. One MA5600T/MA5603T/MA5608T can run only one protocol.
- One MA5600T/MA5603T/MA5608T supports up to eight MG interfaces. If a CKMC daughter board is configured in the MA5600T/MA5603T/MA5608T, up to 32 MG interfaces can be supported. Each MG interface can be configured with the interface attributes independently. Each MG interface can be configured with the attributes (such as authentication and ringing mapping) independently.

Procedure

Configuring the Upstream VLAN Interface

This topic describes how to specify the upstream VLAN interface for the media stream and the signaling flow, and how to configure the IP addresses of the Layer 3 interface. These IP addresses are the sources of the media and signaling IP address pools.

Context

Multiple IP addresses can be configured for the same VLAN Layer 3 interface. Only one IP address functions as the primary address, and other IP addresses function as the secondary addresses.

Procedure

Run the **vlan** command to add an upstream VLAN for the media stream and the signaling flow.

Step 1 Run the **port vlan** command to add the upstream ports to the VLAN.

Step 2 Configure the IP addresses of the VLAN Layer 3 interface.

1. Run the **interface vlanif** command to enter the Layer 3 interface of the upstream VLAN for the media stream and the signaling flow.
2. Run the **ip address** command to configure the IP addresses of the Layer 3 interface.

Step 3 Run the **display interface vlanif** command to check whether the IP addresses of the Layer 3 interface are the same as those in the data plan.

----End

Example

Assume that the media stream and the signaling stream are transmitted upstream through upstream port 0/19/0. To create media and signaling upstream VLAN 3 and configure IP addresses 10.13.4.116/16 and 10.13.4.117/16 of the Layer 3 interfaces for the media IP address pool and the signaling IP address pool respectively, do as follows:

```
huawei(config)#vlan 3 standard
huawei(config)#port vlan 3 0/19 0
huawei(config)#interface vlanif 3
huawei(config-if-vlanif3)#ip address 10.13.4.116 16
huawei(config-if-vlanif3)#ip address 10.13.4.117 16 sub
huawei(config-if-vlanif3)#quit
huawei(config)#display interface vlanif 3
vlanif3 current state : UP
Line protocol current state : UP
Description : HUAWEI, SmartAX Series, vlanif3 Interface
The Maximum Transmit Unit is 1500 bytes
Internet Address is 10.13.4.116/16
Internet Address is 10.13.4.117/16 Secondary
IP Sending Frames' Format is PKTFMT_ETHNT_2, Hardware address is 00e0-fc32-1118
```

Configuring the Media and Signaling IP Address Pools

The media IP address pool defines all media IP addresses that can be used by the AG, and the signaling IP address pool defines all signaling IP addresses that can be used by the AG.

Prerequisites

The IP address of the Layer 3 interface of the media and signaling upstream VLAN must be configured. For details about how to configure the IP address, see Configuring the Upstream VLAN Interface.

Context

- The media IP address and the signaling IP address for the MG interface must be selected from the IP address pools configured here.
- The signaling IP address pool is used to store the IP addresses of the MG interfaces, and the media IP address pool is used to store the IP addresses of the media streams controlled by the signaling.
- The media IP address pool and the signaling IP address pool can be the same. Similarly, the media IP address and the signaling IP address can be the same.

NOTICE

The MGCP interface on the MA5600T/MA5603T/MA5608T does not support the isolation of media streams and signaling flows. Therefore, when the MGCP protocol is used, the media IP address pool and the signaling IP address pool must be the same. When the H.248 or SIP protocol is used, the media IP address pool and the signaling IP address pool can be the same or different.

Procedure

Run the **voip** command to enter the VoIP mode.

Step 1 Configure the media IP address pool.

1. Run the **ip address media** command to add the media IP address to the media IP address pool.

   The media IP address needs to be selected from the IP addresses of the Layer 3 interface of the media and signaling upstream VLAN.

2. Run the **display ip address media** command to check whether the media IP address pool is the same as that in the data plan.

Step 2 Configure the signaling IP address pool.

1. Run the **ip address signaling** command to add the signaling IP address to the signaling IP address pool.

   The signaling IP address needs to be selected from the IP addresses of the Layer 3 interface of the media and signaling upstream VLAN.

2. Run the **display ip address signaling** command to check whether the signaling IP address pool is the same as that in the data plan.

----End

Example

To add IP addresses 10.13.4.116/16 and 10.13.4.117/16 of Layer 3 interfaces of the media and signaling upstream VLAN to the media IP address pool and the signaling IP address pool respectively, do as follows:

```
huawei(config)#voip
huawei(config-voip)#ip address media 10.13.4.116 10.13.0.1
huawei(config-voip)#ip address media 10.13.4.117 10.13.0.1
huawei(config-voip)#ip address signaling 10.13.4.116
huawei(config-voip)#ip address signaling 10.13.4.117
huawei(config-voip)#display ip address media
  Media:
  IP Address...........: 10.13.4.116
  Subnet Mask..........: 255.255.0.0
  Gateway..............: 10.13.0.1
  MAC Address..........: 00-E0-FC-AF-91-33

  IP Address...........: 10.13.4.117
  Subnet Mask..........: 255.255.0.0
  Gateway..............: 10.13.0.1
```

```
 MAC Address..........: 00-E0-FC-AF-91-33
huawei(config-voip)#display ip address signaling
 Signaling:
 IP Address...........: 10.13.4.116
 Subnet Mask..........: 255.255.0.0
 MAC Address..........: 00-E0-FC-AF-91-33

 IP Address...........: 10.13.4.117
 Subnet Mask..........: 255.255.0.0
 MAC Address..........: 00-E0-FC-AF-91-33
```

Adding an MG Interface

This topic describes how to add an MG interface, through which the Access node can communicate with the MGC.

Context

- One Access node supports a maximum of 8 MG interfaces. If a CKMC daughter board is configured in the Access node, up to 32 MG interfaces can be supported. Each MG interface can be configured with the interface attributes independently.
- The configuration of the attributes of an MG interface is valid only to the MG interface.

Procedure

- Add an MG interface that supports H.248.
  a. Run the **display protocol support** command to query the current system protocol.
     - If the current system protocol is H.248, go to h.
     - If the current system protocol is MGCP, go to b.
  b. Run the **display if-mgcp all** command to query whether an MG interface that supports MGCP exists.
     - If such an MG interface does not exist, go to e.
     - If such an MG interface exists, go to c.
  c. Delete all configuration data of this MG interface, and then run the **shutdown(mgcp)** command to disable the MG interface.

NOTICE

This operation directly interrupts all the services on the MG interface. Hence, exercise caution when performing this operation.

  d. Run the **undo interface mgcp** command to delete the MG interface.
  e. Run the **protocol support** command to change the system protocol to H.248.
  f. Run the **save** command to save the configuration data, and then run the **reboot system** command to restart the system to make the new configuration data take effect.
  g. After the system is restarted, log in to the system, and enter the global config mode.
  h. Run the **interface h248** command to add an MG interface that supports H.248.

i. Run the **if-h248 attribute** command to configure the attributes of the MG interface according to the data plan.

j. Run the **display if-h248 attribute** command to check whether the attributes of the MG interface are the same as those in the data plan.

- Add an MG interface that supports MGCP.

a. Run the **display protocol support** command to query the current system protocol.

- If the current system protocol is MGCP, go to h.
- If the current system protocol is H.248, go to b.

b. Run the **display if-h248 all** command to query whether an MG interface that supports H.248 exists.

- If such an MG interface does not exist, go to e.
- If such an MG interface exists, go to c.

c. Delete all configuration data of this MG interface, and then run the **shutdown(h248)** command to disable the MG interface.

CAUTION

This operation directly interrupts all the services on the MG interface. Hence, exercise caution when performing this operation.

d. Run the **undo interface h248** command to delete the MG interface.

e. Run the **protocol support** command to change the system protocol to MGCP.

f. Run the **save** command to save the configuration data, and then run the **reboot system** command to restart the system to make the new configuration data take effect.

g. After the system is restarted, log in to the system, and enter the global config mode.

h. Run the **interface mgcp** command to add an MG interface that supports MGCP.

i. Run the **if-mgcp attribute** command to configure the attributes of the MG interface according to the data plan.

j. Run the **display if-mgcp attribute** command to check whether the attributes of the MG interface are the same as those in the data plan.

----End

Example

Assume that the MG interface ID is 0 and the H.248 protocol is used for interconnecting with the MGC. To add the MG interface, do as follows:

```
huawei(config)#display protocol support
System support H248 protocol
huawei(config)#interface h248 0
  Are you sure to add MG interface?(y/n)[n]:y
```

Assume that the MG interface ID is 0 and the MGCP protocol is used for interconnecting with the MGC. The current system protocol, however, is the H.248 protocol. It is confirmed that

the H.248 interface exists in the system but is not in use. To add MG interface 0, do as follows:

```
huawei(config)#display protocol support
System support H248 protocol
huawei(config)#display if-h248 all
  ---------------------------------------------------------------------------
  MGID    TransMode State    MGPort MGIP          MGCPort MGCIP/DomainName
  ---------------------------------------------------------------------------
  0       -       Close    -     -             -      -
  ---------------------------------------------------------------------------
huawei(config)#undo interface h248 0
  Are you sure to del MG interface?(y/n)[n]:y
huawei(config)#protocol support mgcp
huawei(config)#save
huawei(config)#reboot system
  Please check whether data has saved, the unsaved data will lose if reboot
system, are you sure to reboot system? (y/n)[n]:y
```

After the system is restarted, re-log in to the system.

```
huawei(config)#display protocol support
System support MGCP protocol
huawei(config)#interface mgcp 0
  Are you sure to add MG interface?(y/n)[n]:y
```

(Optional) Configuring the Software Parameters of an MG Interface

The software parameters of an MG interface mainly define certain common service attributes of the MG interface. After the configuration of the software parameters of an MG interface, the software parameters take effect immediately and the configuration is valid only to the MG interface.

Context

There are 34 software parameters (numbered from 0 to 33) of an MG interface that supports H.248. Table 23-43 lists the configurable parameters, and the other parameters are reserved in the system.

Table 23-43 Software parameters of an MG interface that supports H.248

Parameter	Description	Default Setting
2	Indicates whether the MG interface supports dual homing. To configure an MG interface to or not to support dual homing, use this parameter. If the MG interface does not support dual homing (value: 0), even if the secondary MGC is configured, the MG interface does not switch to	0: indicates that dual homing is not supported.

Parameter	Description	Default Setting
	registering with the secondary MGC when the MG interface fails to register with the primary MGC. If the MG interface supports dual homing and auto-switching (value: 2), even if the MG interface has registered with the secondary MGC, the MG interface can automatically switch back to the primary MGC when the primary MGC recovers.	
4	Indicates whether a wildcard is used for the registration of the MG interface. To configure whether a wildcard is used for the registration of an MG interface, use this parameter. Using a wildcard for registration can reduce the quantity of messages between the MG interface and the MGC. When the wildcard is not used for registration, all the terminals on the MG interface need to register with the MGC in order. The registration without a wildcard is also called "single-endpoint registration".	0: indicates that a wildcard is used.
6	Indicates whether the MG interface supports device authentication. To configure an MG interface to or not to support authentication, use this parameter. After the device authentication is supported, run the **auth(h248)** command to configure the authentication parameters, and then run the **reset(h248)** command to reset the MG interface. In this way, the MGC can manage the security of the MGs and avoid illegal registration with the MGC.	1: indicates that device authentication is not supported.

Parameter	Description	Default Setting
7	Indicates whether the MG interface supports security header. To configure an MG interface to or not to support security header, use this parameter.	1: indicates that security header is not supported.
11	Indicates whether the MG interface supports emergency standalone. To configure whether an MG interface supports emergency standalone, use this parameter. If the MG interface supports emergency standalone, the users on the MG interface can make phone calls even if the MG fails to communicate with the MGC.	0: indicates that no call is permitted.
13	Digitmap matching mode	2: indicates the minimum matching.
15	Indicates whether the function of filtering media streams by source port is enabled on an MG interface. To enable or disable the function of filtering media streams by source port on an MG interface, use this parameter. When the function of filtering media streams by source port is enabled on the MG interface, only the media streams from the specific ports can be received.	0: indicates that media streams are not filtered by source port.
16	Indicates the length of the timer for filtering the media stream source port of the MG interface. To configure the length of the timer for filtering the media stream source port of an MG interface, use this parameter. When an MG interface does not filter the source port, the MG interface automatically filters the source port if the	5s

Parameter	Description	Default Setting
	filtering timer times out.	
22	Indicates the type of the prompt tone that is played after the communication between the MG and the MGC is interrupted. To configure the type of the prompt tone after the communication between the MG and the MGC is interrupted, use this parameter.	0: indicates the busy tone.
23	Indicates the length of the timer for synchronizing the port status. To configure the length of the timer for synchronizing the port status, use this parameter.	35s
24	Indicates the maximum value of the Real-Time Transport Protocol (RTP) termination ID.	-
25	Indicates the maximum random value for the protection against avalanche of the H.248 interface.	-
26	Indicates the type of local blocking play tone.	0: indicates the busy tone.
27	Indicates the type of remote blocking play tone.	0: indicates the busy tone.
28	Indicates the duration of the howler tone.	60s
29	Indicates the duration of message waiting tone.	3s
30	Indicates the time limit of the alarm for extra long call.	60 minutes
31	Indicates whether to report the alarm for extra long call.	1: indicates that the alarm is not reported.
32	Min. auto registration interval of remotely-blocked port(s).	1800s
33	Whether MG heartbeat is shut down.	1: No, heartbeat is enabled

There are 17 software parameters (numbered from 0 to 16) of an MG interface that supports MGCP. Table 23-44 lists the configurable parameters, and the other parameters are reserved in the system.

Table 23-44 Software parameters of an MG interface that supports MGCP

Parameter	Description	Default Setting
1	Indicates whether the ongoing call is maintained if the communication between the MG interface and the MGC is abnormal. To maintain the ongoing call when the communication between the MG interface and the MGC is abnormal, use this parameter.	1: disconnects all the calls at once.
2	Indicates whether the MG interface supports dual homing. To configure whether an MG interface supports dual homing, use this parameter. If the MG interface does not support dual homing (value: 1), even if the secondary MGC is configured, the MG interface does not switch to register with the secondary MGC when the MG interface fails to register with the primary MGC.	0: indicates that dual homing is supported.
3	Indicates whether the heartbeat message between the MG and the MGC is disabled. To configure whether the heartbeat message between the MG and the MGC is disabled, use this parameter.	1: indicates that the heartbeat message is not disabled.
4	Indicates whether a wildcard is used for the registration of the MG interface. To configure whether a wildcard is used for the registration of an MG interface, use this parameter. Using a wildcard for registration can reduce the quantity of messages between the MG interface and the	0: indicates that a wildcard is used.

Parameter	Description	Default Setting
	MGC. When a wildcard is not used for registration, all the terminals on the MG interface need to register with the MGC in order. The registration without a wildcard is also called "single-endpoint registration".	
5	Indicates the MGC type. To select the MGC of a different type, use this parameter.	0
6	Indicates the maximum time threshold for responding to the heartbeat messages. To configure the maximum times for transmitting the heartbeat message continuously, use this parameter.	3
7	Indicates whether to report the heartbeat with the MG as an endpoint. To configure whether to report the heartbeat with the MG as an endpoint, use this parameter.	0: indicates that reporting the heartbeat with the MG as an endpoint is not supported.
10	Indicates the point-to-point (P2P) fault reporting. To configure whether the MG interface supports the P2P fault reporting from the ISDN terminals, use this parameter.	0: indicates that the P2P fault is reported.
11	Indicates the point-to-multipoint (P2MP) fault reporting. To configure whether the MG interface supports the P2MP fault reporting from the ISDN terminals, use this parameter.	1: indicates that the P2MP fault is not reported.
12	Indicates the type of local blocking play tone.	0: indicates the busy tone.
13	Indicates the type of remote blocking play tone.	0: indicates the busy tone.
14	Indicates the RTP filtering	1: indicates that the RTP

Parameter	Description	Default Setting
	switch of the MGCP interface. To configure whether the RTP filtering function is enabled, use this parameter. When the RTP filtering function is enabled, only the media stream from the specific ports can be received.	filtering function is not enabled.
15	Indicates the duration of the howler tone.	60s
16	Whether the timer symbol "T" follows the number string reported by the signaling.	0: Yes

Procedure

- When the system protocol is H.248, perform the following operations to configure the software parameters of an MG interface.
  a. In the global config mode, run the **interface h248** command to enter the MG interface mode.
  b. Run the **mg-software parameter** command to configure the software parameters listed in the data plan.
  c. Run the **display mg-software parameter** command to check whether the software parameters of the MG interface are the same as those in the data plan.
- When the system protocol is MGCP, perform the following operations to configure the software parameters of an MG interface.
  a. In the global config mode, run the **interface mgcp** command to enter the MG interface mode.
  b. Run the **mg-software parameter** command to configure the software parameters listed in the data plan.
  c. Run the **display mg-software parameter** command to check whether the software parameters of the MG interface are the same as those in the data plan.

----End

Example

To configure software parameter 11 of H.248-based MG interface 0 to 1 so that the MG interface supports emergency standalone and allows only internal calls, do as follows:

```
huawei(config)#interface h248 0
huawei(config-if-h248-0)#mg-software parameter 11 1
huawei(config-if-h248-0)#display mg-software parameter 11
  -------------------------------------------------
  Interface Id:0         para index:11  value:1
  -------------------------------------------------
 APPENDIX:
```

```
------------------------------------------------
 Interface software parameter name:
 11: Stand alone support
    0: None
    1: Inner
    2: Emergency
    3: Both
```

(Optional) Configuring the TID Format of an MG Interface

The TID format determines how various user terminal IDs on the MG interface are generated.

Prerequisites

NOTICE

The configuration of the TID format is relatively complicated. The information such as the syntax rules with which the terminal prefix must comply and the requirements for the character string of the TID template is defined in the protocol, and is not described here. This topic describes only some basic information. It is recommended that you refer to the TID description in ITU-T H248.1 (applicable to H.248) or RFC 3435 (applicable to MGCP) before configuring the TID format.

Context

The TID format consists of the terminal prefix and the TID template. The TID template defines the generation mode of the TID excluding the terminal prefix. A TID consists of a terminal prefix and a character string generated by a TID template.

The TID templates that are bound to various types of users on the MG interface determine whether the users support terminal layering.

- If the parameter list of the TID template includes only keyword "G", the TID template is used by the non-layering users. Users bound with this template do not support terminal layering.
- If the parameter list of the TID template includes only keywords "F", "S", "P", "B" ("B" is not available to PSTN users), the TID template is used by the layering users. Users bound with this template support terminal layering.

NOTE

The meaning of each keyword is as follows:

- F indicates the subrack ID.
- S indicates the slot ID.
- P indicates the port ID.
- B indicates the B channel ID (only for ISDN BRA and ISDN PRA terminals).
- G indicates the general permanent termination index.
- R indicates the RTP ephemeral termination index (only for the RTP ephemeral termination, which exists only when the system protocol is H.248. This termination is not involved unless special remarks are provided.)

When adding a user that supports terminal layering, you cannot and do not have to specify the parameter **terminalid** because the system automatically allocates a terminal ID. When adding a user that does not support terminal layering, you must specify the parameter **terminalid**.

You can run the **display tid-format(h248)** command or **display tid-format(mgcp)** command to query the TID formats of various types of users on the current MG interface. In the query result, **template-index** indicates the index of the TID template that is bound to the type of users. Then, run the **display tid-template** command to check whether the TID template supports the layering configuration. Hence, you can check whether the user supports terminal layering.

Precaution

- There are 19 default TID templates (templates 0-18) in the system. The default TID templates can be referenced, but cannot be added, modified, or deleted.
- The configuration of terminal layering on the MG must be the same as that on the MGC.
- If certain type of terminals exists on an interface and the interface is not disabled, the terminal prefix of this type of terminals cannot be modified.
- If a certain type of terminals exists on an interface, the TID format (including the terminal prefix and the index of the TID template) of this type of terminals cannot be changed.
- The terminal prefix must comply with the following syntax rules: The prefix must not exceed 64 characters. Only letters, digits, "_", and "/" are the characters allowed for input. The first character must be a letter, and the last character must not be a digit.
- The length of the TID, which is generated by combining the TID template and the terminal prefix that you configured, must not exceed 64 characters.

Procedure

- If the system protocol is H.248, to configure the TID format on the current MG interface, do as follows:
  a. Run the **display tid-template** command to query the information about the default TID template of the system.
  b. If the default TID template cannot meet the service requirements, run the **tid-template add** command to add a TID template that meets the service requirements. If the default TID template meets the service requirements, proceed to c.
  c. Run the **interface h248** command to enter the H.248 mode.
  d. Configure the TID templates and the terminal prefix of various types of users on the current MG (VAG).

NOTICE

The terminal prefix of PSTN, BRA and PRA must be all identical or unique.

  - In the H.248 mode, run the **tid-format rtp** command to configure the TID template and the terminal prefix of the RTP ephemeral termination.
  - In the H.248 mode, run the **tid-format pstn** command to configure the TID template and the terminal prefix of the PSTN user.

- In the H.248 mode, run the **tid-format bra** command to configure the TID template and the terminal prefix of the ISDN BRA user.
- In the H.248 mode, run the **tid-format pra** command to configure the TID template and the terminal prefix of the ISDN PRA user.

e. Run the **display tid-format(h248)** command to check whether the TID templates and the terminal prefixes of various types of users on the current MG interface are the same as those in the data plan.

- If the system protocol is MGCP, to configure the TID format on the current MG interface, do as follows:
  a. Run the **display tid-template** command to query the information about the default TID template of the system.
  b. If the default TID template cannot meet the service requirements, run the **tid-template add** command to add a TID template that meets the service requirements. If the default TID template meets the service requirements, proceed to c.
  c. Run the **interface mgcp** command to enter the MGCP mode.
  d. Configure the TID templates and the terminal prefix of various types of users on the current MG (VAG).
     - In the MGCP mode, run the **tid-format pstn** command to configure the TID template and the terminal prefix of the PSTN user.
     - In the MGCP mode, run the **tid-format bra** command to configure the TID template and the terminal prefix of the ISDN BRA user.
     - In the MGCP mode, run the **tid-format pra** command to configure the TID template and the terminal prefix of the ISDN PRA user.
  e. Run the **display tid-format(mgcp)** command to check whether the TID templates and the terminal prefixes of various types of users on the current MG interface are the same as those in the data plan.

----End

Example

Assume that in the H.248 mode, the terminal prefix of the PSTN user on MG interface 1 is aln/, and the layering TID template 3 is used. To add a PSTN user on port 0/2/0 and check whether the system automatically allocates a TID generated according to the template, do as follows:

```
huawei(config)#display tid-template 3//Query the information about TID template 3
  -------------------------------------------------
  Index     : 3
  Format    : %u/%u/%u
  Para-list  : F+1,S+1,P+1  //The parameter list of the TID template includes keyword
"F", "S",
               //and "P", which indicates that this template supports terminal layering.
  Name      : Aln_Not_Fixed_1
  -------------------------------------------------

huawei(config)#interface h248 1
huawei(config-if-h248-1)#tid-format pstn prefix aln template 3
huawei(config-if-h248-1)#quit
huawei(config)#esl user
huawei(config-esl-user)#mgpstnuser add 0/2/0 1
```

```
huawei(config-esl-user)#display mgpstnuser 0/2/0
{ <cr>|endframeid/slotid/portid<S><1,15> }:

 Command:
        display mgpstnuser 0/15/0
 ------------------------------------------------------------------------
 F  /S /P    MGID    TelNo          Priority PotsLineType TerminalID
 ------------------------------------------------------------------------
 0  /2 /0    1       -             Cat3    DEL        aln/1/3/1
          //The system allocates the terminal ID according to the TID format.
 ------------------------------------------------------------------------
```

Enabling an MG Interface

Enabling an MG interface is to reset an MG interface to make the MG interface register with the MGC (or to make the modified attributes of the MG interface take effect) after the configuration of the MG interface is complete, so that the MG interface can work in the normal state.

Precaution

CAUTION

For the MG interface that has been in service, this operation interrupts the ongoing services carried on the MG interface. Hence, exercise caution when performing this operation.

Procedure

- Enable the MG interface that adopts the H.248 protocol.
  a. Run the **interface h248** command to enter the H.248 mode.
  b. Run the **reset coldstart** command to enable the MG interface.
  c. Run the **quit** command to return to the global config mode, and then run the **display if-h248 all** command to check whether the MG interface is in the normal state.
- Enable the MG interface that adopts the MGCP protocol.
  a. Run the **interface mgcp** command to enter the MGCP mode.
  b. Run the **reset** command to enable the MG interface.
  c. Run the **quit** command to return to the global config mode, and then run the **display if-mgcp all** command to check whether the MG interface is in the normal state.

----End

Example

To enable H.248-based MG interface 0, do as follows:

```
huawei(config)#interface h248 0
huawei(config-if-h248-0)#reset coldstart
  Are you sure to reset MG interface?(y/n)[n]:y
huawei(config-if-h248-0)#quit
huawei(config)#display if-h248 all
--------------------------------------------------------------------------------
MGID    TransMode State        MGPort MGIP         MGCPort MGCIP/DomainName
--------------------------------------------------------------------------------
0       UDP      Normal       2944   10.10.10.11    2944   10.10.20.11
--------------------------------------------------------------------------------
```

To enable MGCP-based MG interface 0, do as follows:

```
huawei(config)#interface mgcp 0
huawei(config-if-mgcp-0)#reset
  Are you sure to reset MG interface?(y/n)[n]:y
huawei(config-if-mgcp-0)#quit
huawei(config)#display if-mgcp all
  ------------------------------------------------------------------------------
  MGID    State        MGPort MGIP          MGCPort MGCIP/DomainName
  ------------------------------------------------------------------------------
  0       Normal       2727   10.10.10.11    2727    10.10.20.11
  1       Wait ack      2527   10.10.10.12    2727    10.10.20.12
  ------------------------------------------------------------------------------
```

23.21.2 Configuring the IUA Link

This topic describes how to configure the IUA link for signaling transmission between the Access node and MGC in the VoIP ISDN service.

Context

- Simple Control Transmission Protocol (SCTP) is a connection-oriented protocol. Its most fundamental function is to provide reliable transmission for interaction messages between the Access node and MGC. The SCTP protocol implements services based on the association between two SCTP endpoints. SCTP can be regarded as a transmission layer. Its upper layer is called SCTP subscriber, and its lower layer is the IP network.
- The IUA link is the carrier of the interaction signaling between the Access node and MGC.

Adding an IUA Link Set

This topic describes how to add an IUA link set. When configuring the VoIP ISDN service, you need to configure the IUA link to carry the Q.931 call signaling. Before adding an IUA link, you must add a corresponding link set. Otherwise, the link cannot be added.

Context

- The system supports the configuration of a maximum of IUA link sets.
- After a link set is configured successfully, it is in the deactivated state by default.

Procedure

In global config mode, run the command**sigtran** to enter the Sigtran mode.

Step 1 Run the **iua-linkset add** command to add an IUA link set.

Step 2 You can run the **display iua-linkset attribute** command to check whether the configured IUA link set information is the same as the data plan.

----End

Example

Assume that the link set ID is 0, working mode of the link set is active/standby mode, pending duration is 20s, prefix of the IID is b/, IID generation mode is using the binary value that is automatically generated in ffsspp mode, namely, parameter 2. To add such an IUA link set, do as follows:

```
huawei(config)#sigtran
```

Adding an IUA Link

This topic describes how to add an IUA link. After the link set is added, you can add an IUA link to carry the Q.931 call signaling for the ISDN user.

Prerequisites

The IUA link set must be added.

Context

- Make sure that a minimum of one item in the local port ID, local IP address, remote port ID, and remote IP address of a link is different from the corresponding item of other links.
- Only two links can be configured in the same link set. In addition, the local IP addresses of the two links must be the same.

Procedure

(This step is not required if the command line interface is already in the Sigtran mode.) In global config mode, run the **sigtran** command to enter the Sigtran mode.

Step 1 Run the **iua-link add** command to add an IUA link.

Step 2 You can run the **display iua-link attribute** command to check whether the configured IUA link information is the same as the data plan.

----End

Example

Assume that the link ID is 0, link set ID is 0, local port ID is 1402, local IP address is 10.10.10.10, remote port ID is 1404, and remote IP address 1 is 10.10.10.20. To add such a link, do as follows:

```
huawei(config)#sigtran
huawei(config-sigtran)#iua-link add 0 0 1402 10.10.10.10 1404 10.10.10.20
huawei(config-sigtran)#display iua-link attribute
{ <cr>|linkno<L> }:

 Command:
       display iua-link attribute
 ---------------------------------------------------------------------------
 LinkNo             :  0
 LinksetNo          :  0
 Local port         :  1402
 Local IP address    :  10.10.10.10
 Remote port        :  1404
 Remote IP address   :  10.10.10.20
 Remote IP address 2  :  -
 Priority           :  0
 ---------------------------------------------------------------------------
```

23.21.3 Configuring the VoIP ISDN BRA User

After the MG interface is configured, you can add the VoIP ISDN BRA user on this interface, and configure the system parameters, oversea parameters, and attributes of the BRA port, to implement the VoIP ISDN BRA service.

Configuring the ISDN BRA User Data

This topic describes how to configure the ISDN BRA user data based on H.248 (including the priority, flag of reporting the fault of UNI, etc., and the data must be the same as the corresponding data on the MGC) so that the ISDN BRA user can access the network to use the ISDN BRA service.

Prerequisites

The ISDN BRA service board must be inserted into the planned slot correctly and the board must be in the normal state.

The IUA link must be configured according to the requirements. For details about how to configure the IUA link, see 23.21.2 Configuring the IUA Link.

Default Configuration

Table 23-45 lists the default settings of the attributes of the ISDN BRA user. When configuring these attributes, you can modify the values according to the service requirements.

Table 23-45 Default settings of the attributes of the ISDN BRA user

Parameter	Default Settings
Priority of the ISDN BRA user	cat3: (common user)
Flag of reporting the UNI fault of the ISDN BRA user	Disable
Threshold for the number of auto recoveries	20

Parameter	Default Settings
from deterioration faults	

Procedure

In global config mode, run the **esl user** command to enter the ESL user mode.

Step 1 Run the **mgbrauser add** command to add an ISDN BRA.

- When **iid-map** in the **iua-linkset add** command is configured to 1, interfaceid must be configured and be different from the interfaceid of other users in the same link set.
- The terminal ID of an ISDN BRA user must be different from the terminal IDs of other users.
- If the MG interface does not support the terminal layering function, the terminal ID must be configured when an ISDN BRA user is added. In addition, the terminal ID must differ from the terminal ID of the existing ISDN BRA user by an integer multiple of 2. For example, to add the first ISDN BRA user, the terminal ID is 2; to add the second ISDN BRA user, the terminal ID is 4; to add the third ISDN BRA user, the terminal ID is 6; the rest may be deduced by analogy.
- If the MG interface supports the terminal layering function, the terminal ID cannot be configured when an ISDN BRA user is added on the MG interface. The system automatically allocates a terminal ID for the user.

Step 2 Run the **mgbrauser attribute set** command to configure the attributes of the ISDN BRA user.

The attributes of an ISDN BRA user include the following:

- Priority of the ISDN BRA user. The priorities are classified to cat1 (the first class government user), cat2 (the second class government user), and cat3 (common user) in the sequence of descending. Without special requirements, the default cat3 is adopted.
- Flag of reporting the UNI fault of the ISDN BRA user. This attribute controls whether to report the UNI fault to MGC. The default is not to report.
- Threshold for the number of auto recoveries from deterioration faults. This attribute indicates the maximum times that the equipment attempts to recover from deterioration faults. Zero indicates not to recover automatically. The number of 255 indicates not to limit the attempt times. The default is 20.

Step 3 Run the **display mgbrauser attribute** command to query whether the configured attributes of the ISDN BRA user are the same as the data plan.

----End

Example

Assume the following configurations:

- Link set ID: 0
- IUA interface ID: 0 (value of **iid-map**: 1)
- Terminal ID: 2 (not supporting the terminal layering function)
- User priority: cat3

- Telephone number: 83110001 (Before configuring a emergency standalone number, ensure that the emergency standalone function has been enabled on the MG interface. For details, see (Optional) Configuring the Software Parameters of an MG Interface.)
- UNI alarm report function: enable
- Threshold for the number of auto recoveries from L1 deterioration faults: 30

To add an ISDN BRA user with such configurations on port 0/4/0 of MG interface 0, do as follows:

```
huawei(config-sigtran)#quit
huawei(config)#esl user
huawei(config-esl-user)#mgbrauser add 0/4/0 0 0 interfaceid 0 terminalid 2 priority
cat3 telno 83110001
 Are you sure to configure the working mode of the DSL board to normal and reset
 the board automatically? (y/n)[n]:y

huawei(config-esl-user)#display mgbrauser 0/4
  ---------------------------------------------------------------------------
  F  /S /P /B  MGID    LinkSetNo UserIFID TelNo          Priority TerminalID
  ---------------------------------------------------------------------------
  0/4/0 /0  0       0        0       83110001        Cat3    A2
  ---------------------------------------------------------------------------
huawei(config-esl-user)#mgbrauser attribute set 0/4/0 priority cat3 unireport enable
auto-resume-limit 30
huawei(config-esl-user)#display mgbrauser attribute 0/4/0
{ <cr>|endframeid/slotid/portid<S><Length 1-15> }:

  Command:
         display mgbrauser attribute 0/4/0
 ---------------------------------------------------------------------------
  F  /S /P                      :  0/4/0
  UNIreport                     :  enable
  Priority                      :  Cat3
  Auto reservice times/limit       :  0/30
  DSP-para-template              :  -
  ---------------------------------------------------------------------------
```

(Optional) Configuring the System Parameters

This topic describes how to configure the system parameters including the overseas version flag and message waiting indication (MWI) mode according to the local standards to ensure that the response of the user terminal complies with the local standards.

Procedure

Run the **system parameters** command to configure the system parameters.

Step 1 Run the **display system parameters** command to check whether the system parameters are the same as those in the data plan.

----End

Example

To configure the overseas version flag (system parameter 1) to Hong Kong (parameter value 1), do as follows:

```
huawei(config)#system parameters 1 1
huawei(config)#display system parameters 1
  ---------------------------------------------------------------------------
  Parameter name index: 1    Parameter value: 1
  Mean: Overseas version flag, 0:China, 1:HongKong, 2:Brazil, 3:Egypt, 4:
Singapore, 5:Thailand, 6:France, 7:Britain MSFUK, 8:Britain ETSI, 9:Bulgaria,
10:Reserved, 11:Austria, 12:Hungary, 13:Poland
  ---------------------------------------------------------------------------
```

(Optional) Configuring the Overseas Parameters

By default, the overseas parameters are configured according to the Chinese standards. In the actual service configuration, the attributes such as the upper and lower thresholds of the flash-hooking duration can be configured according to the local standards to ensure that the response of the user terminal complies with the local standards.

Procedure

Run the **oversea parameters** command to configure the overseas parameters.

Step 1 Run the **display oversea parameters** command to check whether the overseas parameters are the same as those in the data plan.

----End

Example

To set the upper flash-hooking threshold (overseas feature parameter 0) to 800 ms (in compliance with the Hong Kong standard) and the lower flash-hooking threshold (overseas feature parameter 1) to 100 ms (in compliance with the Hong Kong standard), do as follows:

```
huawei(config)#oversea parameters 0 800
huawei(config)#oversea parameters 1 100
huawei(config)#display oversea parameters
{ <cr>|name<U><0,99> }:

  Command:
         display oversea parameters
  ---------------------------------------------------------------------------
  Parameter name index: 0    Parameter value: 800
  Mean: Hooking upper threshold(ms), reference: China:350, HongKong:800
  ---------------------------------------------------------------------------
  Parameter name index: 1    Parameter value: 100
  Mean: Hooking lower threshold(ms), reference: China:100, HongKong:100
  ---------------------------------------------------------------------------
  Parameter name index: 2    Parameter value: 0
  Mean: Flag of applying PARKED LINE FEED or not when user port is locked, 0:not
 apply, 1:apply
  ---------------------------------------------------------------------------
  Parameter name index: 3    Parameter value: 0
```

```
  Mean: The detect time of flash upper limit to onhook, default value: 0ms
  ----------------------------------------------------------------------------
  Parameter name index: 4    Parameter value: 0
  Mean: DTMF Detector Tuning, 0:Q.24, 1:Russia
----------------------------------------------------------------------------
  Parameter name index: 5    Parameter value: 0
  Mean: Flag of changing TEI value for software switch, 0:no change, 1:set TEI
value to 1, 2:set TEI value to 0. Default: 0
  ----------------------------------------------------------------------------
```

(Optional) Configuring the Attributes of an ISDN BRA Port

This topic describes how to configure the attributes of an ISDN BRA port to ensure that the ISDN BRA port can meet the actual application requirements. You can configure the auto-deactivation status, remote power supply status, UNI fault alarming function, and working mode of the port.

Default Configuration

Table 23-46 lists the default values of the attributes of an ISDN BRA port. When configuring the attributes, you can change the values according to the service requirements.

Table 23-46 Default values of the attributes of an ISDN BRA port

Parameter	Default Setting
Autodeactive	Disable
Autodeactive-delay	30s
Activemode	unstable-active
Remotepower	Disable
Unialarm	Disable
Workmode	p2mp

Procedure

In global config mode, run the **braport** command to enter braport mode.

Step 1 Run the **braport attribute set** command to configure the attributes such as the working mode, auto-deactivation status, and remote power supply status of the port.

If an ISDN BRA port needs to be connected to multiple terminal users, configure the working mode of the port to p2mp. If an ISDN BRA port needs to be connected to only one terminal user, configure the working mode of the port to p2p.

For detailed description of the **braport attribute set** command, see the parameter description in **braport attribute set**.

----End

Example

Assume that the working mode is p2mp, the activation mode is stable, and the auto-deactivation function is disabled. To configure such attributes of ISDN BRA port 0/4/0, do as follows:

```
huawei(config)#braport
huawei(config-braport)#braport attribute set 0/4/0 workmode p2mp activemode
stable-active
huawei(config-braport)#display braport attribute
{ frameid/slotid/portid<S><Length 1-15>|frameid/slotid<S><Length 1-15> }:0/4/0

  Command:
         display braport attribute 0/4/0
-------------------------------------------------------------------------------
F  /S /P  Remotepower Workmode Autodeactive Deactivedelay Activemode Unialarm
-------------------------------------------------------------------------------
0/4/0  disable    p2mp    disable     30          stable    disable
-------------------------------------------------------------------------------
```

23.22 Configuring the VoIP ISDN PRA Service (H.248-based)

This topic describes how to configure the VoIP ISDN PRA service on an IP network. When the Access node uses the H.248 protocol, the device supports the access of the ISDN PRA user. ISDN provides end-to-end (E2E) digital connection and supports multiple types of voice and non-voice telecom services.

Prerequisites

According to the actual network, a route from the Access node to the MGC must be configured to ensure that the Access node communicates with the MGC normally.

Context

- The voice over IP (VoIP) service uses the IP packet switched network for transmission after the traditional analog voice signals are compressed and packetized. This service lowers the cost of the voice service. For the detailed description of the VoIP service, see 23 Voice Feature in the *Feature Description*.
- Defined by the International Telegraph and Telephone Consultative Committee (CCITT), the ISDN is a communication network evolved from the Integrated Digital Network (IDN). The ISDN service provides the E2E digital connection and supports multiple types of voice and non-voice telecom services. On the ISDN network, users can access the network through the following two interfaces: (The ISDN technology provides two types of user-network interfaces based on the user-network interface reference model.)
  - ISDN basic rate interface (BRI), which supports a rate of 144 kbit/s and provides two B channels (for carrying services) and one D channel (for transmitting call control signaling and maintenance and management signaling). The rate of B channels is 64 kbit/s and the rate of D channel is 16 kbit/s. The service carried on the ISDN BRI is the ISDN basic rate access (BRA) service.

- ISDN primary rate interface (PRI), which supports a rate of 2.048 Mbit/s and provides 30 B channels and one D channel. The rates of the B channel and D channel are both 64 kbit/s. The service carried on the ISDN PRI is the ISDN primary rate access (PRA) service.

Precaution

The media gateway control protocol (MGCP) is a master/slave protocol, under which the MGC controls the AG to implement the call connection. The data on the AG for the interconnection with the MGC, such as the attributes of the MG interface and the attributes of the VoIP user, must be the same as the corresponding data on the MGC. Therefore, before configuring the VoIP service, you must contact MGC engineers to check and ensure that the interconnection data plan for the AG is consistent with the corresponding plan for the MGC.

Data preparation

Table 23-47 provides the data plan for configuring the H.248-based VoIP ISDN PRA service.

Table 23-47 Data plan for configuring the H.248-based VoIP ISDN PRA service

Item			Remarks
MG interface data (The data must be consistent with the data on the MGC.)	Parameters of the media stream and signaling stream	Upstream VLAN for media and signaling streams	It is used as the upstream VLAN of the VoIP service to be configured. Standard VLAN is recommended.
		Uplink port for media and signaling streams	It is used as the uplink port for the VoIP service to be configured.
		Media IP address and signaling IP address	These IP addresses must be contained in the media and signaling IP address pools. The media and signaling IP address pools consist of all the IP addresses of the L3 interface of the upstream VLAN for media and signaling streams.
		Default IP address of the MG	It is the next hop IP address from the Access node to the MGC.
	Attribute parameters of the MG interface **NOTE** Parameters listed here are mandatory, which means that the MG	MG interface ID	It is the ID of the MG interface used by the VoIP service to be configured. The Access node supports only one VAG.
		Signaling port ID of the MG interface	It is the transport layer protocol port ID used for the signaling exchange between the Access node (AG) and the MGC. The default signaling port ID

Item			Remarks
	interface cannot be started if these parameters are not configured.		defined in H.248 is 2944 (text) and 2945 (binary).
		IP address of the primary MGC to which the MG interface belongs	When dual homing is not configured, you can configure the parameters of only the primary MGC. When dual homing is configured, you also need to configure the IP address and port ID of the secondary MGC.
		Port ID of the primary MGC to which the MG interface belongs	
		Coding mode of the MG interface	Currently, only the **text** mode is supported.
		Transmission mode of the MG interface	The transmission mode is selected according to the requirements of the MGC. Generally, UDP is used.
		Domain name of the MG interface	It corresponds to the **domainName** parameter of the MG interface. When the H.248 protocol is used, this parameter must be configured if the **MIDType** parameter of the H.248 message is configured to **domainName**. Otherwise, the MG interface cannot be started. In other situations, this parameter is optional.
		Device name of the MG interface	It is supported by only the H.248 protocol, and it corresponds to the **deviceName** parameter of the MG interface that uses the H.248 protocol. This parameter must be configured if the **MIDType** parameter of the H.248 message is configured to **domainName**. Otherwise, the MG interface cannot be started. In other situations, this parameter is optional.
	(Optional) Configuring the Software Parameters of an MG		Whether the MG interface supports the functions such as dual homing and emergency

Item			Remarks
	Interface		standalone is determined by the service requirements.
	(Optional) Configuring the TID Format of an MG Interface		The TID format determines the generation mode of various types of user terminals on an MG interface.
IUA link	Adding an IUA Link Set		The IUA link can be configured only after the IUA link set is configured.
	Adding an IUA Link **NOTE** The local port ID, local IP address, remote port ID, and remote IP address of different links must not be completely same; otherwise, the service cannot be configured.	IUA link ID	It indicates the link for transmitting the signaling.
		IUA link set ID	-
		Local port ID	To activate the link normally, it must be the same as the remote port ID configured on the MGC.
		Local IP address	It must be the same as the remote IP address of the link configured on the MGC. In addition, the local IP addresses of the links that are in the same link set must be the same. (The IP address must exist in the media IP address pool.)
		Remote port ID	To activate the link normally, it must be the same as the local port ID configured on the MGC.
		Remote IP address	It must be the same as the local IP address of the link configured on the MGC. The SCTP protocol supports the multi-homing function. That is, one link can be configured with the IP addresses of multiple MGCs as the remote IP addresses. When one MGC is faulty, the link can be switched to other MGCs automatically. This ensures that the service is not affected. The Access node supports the configuration of the active remote IP address and standby remote IP address.
ISDN PRA user data (The data must be consistent with the	Slot that houses the E1 service board		-
	Configuring	Termination	If the TID format bound to the

Item			Remarks
data on the MGC.)	the ISDN PRA User Data	ID	PRA user does not support terminal layering function, this parameter needs to be configured, and the configuration must be consistent with the configuration on the MGC.
		User priority	The user priority must be specified according to the service requirements. There are three categories of user priorities, which are as follows: • cat1: government1 (category 1 government user) • cat2: government2 (category 2 government user) • cat3: normal (common user, default)
		Interface ID	It indicates the interface for the PRA user data to pass through the MG and MGC. The configuration of this parameter must be consistent with the corresponding configuration on the MGC.
	(Optional) Configuring the System Parameters		The system parameters including the oversea version flag and message waiting indication (MWI) mode need to be configured according to the local standards to ensure that the response of the user terminal complies with the local standards.
	(Optional) Configuring the Overseas Parameters		The attributes such as the upper and lower thresholds of the flash-hooking duration need to be configured according to the local standards to ensure that the response of the user terminal complies with the local standards.
	Configuring the Attributes of the E1 Port		The attributes include the access mode of the board, port mode, line coding mode, and port impedance. Generally, if there is no requirement, these attributes need not be modified.

23.22.1 Configuring an MG Interface

This topic describes how to configure an MG interface for implementing the communication between the MA5600T/MA5603T/MA5608T (AG) and the MGC.

Context

- The MA5600T/MA5603T/MA5608T communicates with the MGC through either the H.248 or the MGCP protocol. One MA5600T/MA5603T/MA5608T can run only one protocol.
- One MA5600T/MA5603T/MA5608T supports up to eight MG interfaces. If a CKMC daughter board is configured in the MA5600T/MA5603T/MA5608T, up to 32 MG interfaces can be supported. Each MG interface can be configured with the interface attributes independently. Each MG interface can be configured with the attributes (such as authentication and ringing mapping) independently.

Procedure

Configuring the Upstream VLAN Interface

This topic describes how to specify the upstream VLAN interface for the media stream and the signaling flow, and how to configure the IP addresses of the Layer 3 interface. These IP addresses are the sources of the media and signaling IP address pools.

Context

Multiple IP addresses can be configured for the same VLAN Layer 3 interface. Only one IP address functions as the primary address, and other IP addresses function as the secondary addresses.

Procedure

Run the **vlan** command to add an upstream VLAN for the media stream and the signaling flow.

Step 1 Run the **port vlan** command to add the upstream ports to the VLAN.

Step 2 Configure the IP addresses of the VLAN Layer 3 interface.

1. Run the **interface vlanif** command to enter the Layer 3 interface of the upstream VLAN for the media stream and the signaling flow.
2. Run the **ip address** command to configure the IP addresses of the Layer 3 interface.

Step 3 Run the **display interface vlanif** command to check whether the IP addresses of the Layer 3 interface are the same as those in the data plan.

----End

Example

Assume that the media stream and the signaling stream are transmitted upstream through upstream port 0/19/0. To create media and signaling upstream VLAN 3 and configure IP

addresses 10.13.4.116/16 and 10.13.4.117/16 of the Layer 3 interfaces for the media IP address pool and the signaling IP address pool respectively, do as follows:

```
huawei(config)#vlan 3 standard
huawei(config)#port vlan 3 0/19 0
huawei(config)#interface vlanif 3
huawei(config-if-vlanif3)#ip address 10.13.4.116 16
huawei(config-if-vlanif3)#ip address 10.13.4.117 16 sub
huawei(config-if-vlanif3)#quit
huawei(config)#display interface vlanif 3
vlanif3 current state : UP
Line protocol current state : UP
Description : HUAWEI, SmartAX Series, vlanif3 Interface
The Maximum Transmit Unit is 1500 bytes
Internet Address is 10.13.4.116/16
Internet Address is 10.13.4.117/16 Secondary
IP Sending Frames' Format is PKTFMT_ETHNT_2, Hardware address is 00e0-fc32-1118
```

Configuring the Media and Signaling IP Address Pools

The media IP address pool defines all media IP addresses that can be used by the AG, and the signaling IP address pool defines all signaling IP addresses that can be used by the AG.

Prerequisites

The IP address of the Layer 3 interface of the media and signaling upstream VLAN must be configured. For details about how to configure the IP address, see Configuring the Upstream VLAN Interface.

Context

- The media IP address and the signaling IP address for the MG interface must be selected from the IP address pools configured here.
- The signaling IP address pool is used to store the IP addresses of the MG interfaces, and the media IP address pool is used to store the IP addresses of the media streams controlled by the signaling.
- The media IP address pool and the signaling IP address pool can be the same. Similarly, the media IP address and the signaling IP address can be the same.

NOTICE

The MGCP interface on the MA5600T/MA5603T/MA5608T does not support the isolation of media streams and signaling flows. Therefore, when the MGCP protocol is used, the media IP address pool and the signaling IP address pool must be the same. When the H.248 or SIP protocol is used, the media IP address pool and the signaling IP address pool can be the same or different.

Procedure

Run the **voip** command to enter the VoIP mode.

Step 1 Configure the media IP address pool.

1. Run the **ip address media** command to add the media IP address to the media IP address pool.

   The media IP address needs to be selected from the IP addresses of the Layer 3 interface of the media and signaling upstream VLAN.

2. Run the **display ip address media** command to check whether the media IP address pool is the same as that in the data plan.

Step 2 Configure the signaling IP address pool.

1. Run the **ip address signaling** command to add the signaling IP address to the signaling IP address pool.

   The signaling IP address needs to be selected from the IP addresses of the Layer 3 interface of the media and signaling upstream VLAN.

2. Run the **display ip address signaling** command to check whether the signaling IP address pool is the same as that in the data plan.

----End

Example

To add IP addresses 10.13.4.116/16 and 10.13.4.117/16 of Layer 3 interfaces of the media and signaling upstream VLAN to the media IP address pool and the signaling IP address pool respectively, do as follows:

```
huawei(config)#voip
huawei(config-voip)#ip address media 10.13.4.116 10.13.0.1
huawei(config-voip)#ip address media 10.13.4.117 10.13.0.1
huawei(config-voip)#ip address signaling 10.13.4.116
huawei(config-voip)#ip address signaling 10.13.4.117
huawei(config-voip)#display ip address media
  Media:
  IP Address...........: 10.13.4.116
  Subnet Mask..........: 255.255.0.0
  Gateway..............: 10.13.0.1
  MAC Address..........: 00-E0-FC-AF-91-33

  IP Address...........: 10.13.4.117
  Subnet Mask..........: 255.255.0.0
  Gateway..............: 10.13.0.1
  MAC Address..........: 00-E0-FC-AF-91-33
huawei(config-voip)#display ip address signaling
  Signaling:
  IP Address...........: 10.13.4.116
  Subnet Mask..........: 255.255.0.0
  MAC Address..........: 00-E0-FC-AF-91-33

  IP Address...........: 10.13.4.117
  Subnet Mask..........: 255.255.0.0
  MAC Address..........: 00-E0-FC-AF-91-33
```

Adding an MG Interface

This topic describes how to add an MG interface, through which the Access node can communicate with the MGC.

Context

- One Access node supports a maximum of 8 MG interfaces. If a CKMC daughter board is configured in the Access node, up to 32 MG interfaces can be supported. Each MG interface can be configured with the interface attributes independently.
- The configuration of the attributes of an MG interface is valid only to the MG interface.

Procedure

- Add an MG interface that supports H.248.
  a. Run the **display protocol support** command to query the current system protocol.
     - If the current system protocol is H.248, go to h.
     - If the current system protocol is MGCP, go to b.
  b. Run the **display if-mgcp all** command to query whether an MG interface that supports MGCP exists.
     - If such an MG interface does not exist, go to e.
     - If such an MG interface exists, go to c.
  c. Delete all configuration data of this MG interface, and then run the **shutdown(mgcp)** command to disable the MG interface.

NOTICE

This operation directly interrupts all the services on the MG interface. Hence, exercise caution when performing this operation.

  d. Run the **undo interface mgcp** command to delete the MG interface.
  e. Run the **protocol support** command to change the system protocol to H.248.
  f. Run the **save** command to save the configuration data, and then run the **reboot system** command to restart the system to make the new configuration data take effect.
  g. After the system is restarted, log in to the system, and enter the global config mode.
  h. Run the **interface h248** command to add an MG interface that supports H.248.
  i. Run the **if-h248 attribute** command to configure the attributes of the MG interface according to the data plan.
  j. Run the **display if-h248 attribute** command to check whether the attributes of the MG interface are the same as those in the data plan.
- Add an MG interface that supports MGCP.
  a. Run the **display protocol support** command to query the current system protocol.
     - If the current system protocol is MGCP, go to h.
     - If the current system protocol is H.248, go to b.
  b. Run the **display if-h248 all** command to query whether an MG interface that supports H.248 exists.

- If such an MG interface does not exist, go to e.
- If such an MG interface exists, go to c.

c. Delete all configuration data of this MG interface, and then run the **shutdown(h248)** command to disable the MG interface.

CAUTION

This operation directly interrupts all the services on the MG interface. Hence, exercise caution when performing this operation.

d. Run the **undo interface h248** command to delete the MG interface.

e. Run the **protocol support** command to change the system protocol to MGCP.

f. Run the **save** command to save the configuration data, and then run the **reboot system** command to restart the system to make the new configuration data take effect.

g. After the system is restarted, log in to the system, and enter the global config mode.

h. Run the **interface mgcp** command to add an MG interface that supports MGCP.

i. Run the **if-mgcp attribute** command to configure the attributes of the MG interface according to the data plan.

j. Run the **display if-mgcp attribute** command to check whether the attributes of the MG interface are the same as those in the data plan.

----End

Example

Assume that the MG interface ID is 0 and the H.248 protocol is used for interconnecting with the MGC. To add the MG interface, do as follows:

```
huawei(config)#display protocol support
System support H248 protocol
huawei(config)#interface h248 0
  Are you sure to add MG interface?(y/n)[n]:y
```

Assume that the MG interface ID is 0 and the MGCP protocol is used for interconnecting with the MGC. The current system protocol, however, is the H.248 protocol. It is confirmed that the H.248 interface exists in the system but is not in use. To add MG interface 0, do as follows:

```
huawei(config)#display protocol support
System support H248 protocol
huawei(config)#display if-h248 all
  ------------------------------------------------------------------------------
  MGID    TransMode State    MGPort MGIP          MGCPort MGCIP/DomainName
  ------------------------------------------------------------------------------
  0       -       Close    -     -             -      -
  ------------------------------------------------------------------------------
huawei(config)#undo interface h248 0
  Are you sure to del MG interface?(y/n)[n]:y
huawei(config)#protocol support mgcp
```

```
huawei(config)#save
huawei(config)#reboot system
  Please check whether data has saved, the unsaved data will lose if reboot
system, are you sure to reboot system? (y/n)[n]:y
```

After the system is restarted, re-log in to the system.

```
huawei(config)#display protocol support
System support MGCP protocol
huawei(config)#interface mgcp 0
  Are you sure to add MG interface?(y/n)[n]:y
```

(Optional) Configuring the Software Parameters of an MG Interface

The software parameters of an MG interface mainly define certain common service attributes of the MG interface. After the configuration of the software parameters of an MG interface, the software parameters take effect immediately and the configuration is valid only to the MG interface.

Context

There are 34 software parameters (numbered from 0 to 33) of an MG interface that supports H.248. Table 23-48 lists the configurable parameters, and the other parameters are reserved in the system.

Table 23-48 Software parameters of an MG interface that supports H.248

Parameter	Description	Default Setting
2	Indicates whether the MG interface supports dual homing. To configure an MG interface to or not to support dual homing, use this parameter. If the MG interface does not support dual homing (value: 0), even if the secondary MGC is configured, the MG interface does not switch to registering with the secondary MGC when the MG interface fails to register with the primary MGC. If the MG interface supports dual homing and auto-switching (value: 2), even if the MG interface has registered with the secondary MGC, the MG interface can automatically switch back to the primary MGC when the primary MGC recovers.	0: indicates that dual homing is not supported.
4	Indicates whether a wildcard is	0: indicates that a wildcard is

Parameter	Description	Default Setting
	used for the registration of the MG interface. To configure whether a wildcard is used for the registration of an MG interface, use this parameter. Using a wildcard for registration can reduce the quantity of messages between the MG interface and the MGC. When the wildcard is not used for registration, all the terminals on the MG interface need to register with the MGC in order. The registration without a wildcard is also called "single-endpoint registration".	used.
6	Indicates whether the MG interface supports device authentication. To configure an MG interface to or not to support authentication, use this parameter. After the device authentication is supported, run the **auth(h248)** command to configure the authentication parameters, and then run the **reset(h248)** command to reset the MG interface. In this way, the MGC can manage the security of the MGs and avoid illegal registration with the MGC.	1: indicates that device authentication is not supported.
7	Indicates whether the MG interface supports security header. To configure an MG interface to or not to support security header, use this parameter.	1: indicates that security header is not supported.
11	Indicates whether the MG interface supports emergency standalone. To configure whether an MG interface supports emergency standalone, use this parameter.	0: indicates that no call is permitted.

Parameter	Description	Default Setting
	If the MG interface supports emergency standalone, the users on the MG interface can make phone calls even if the MG fails to communicate with the MGC.	
13	Digitmap matching mode	2: indicates the minimum matching.
15	Indicates whether the function of filtering media streams by source port is enabled on an MG interface. To enable or disable the function of filtering media streams by source port on an MG interface, use this parameter. When the function of filtering media streams by source port is enabled on the MG interface, only the media streams from the specific ports can be received.	0: indicates that media streams are not filtered by source port.
16	Indicates the length of the timer for filtering the media stream source port of the MG interface. To configure the length of the timer for filtering the media stream source port of an MG interface, use this parameter. When an MG interface does not filter the source port, the MG interface automatically filters the source port if the filtering timer times out.	5s
22	Indicates the type of the prompt tone that is played after the communication between the MG and the MGC is interrupted. To configure the type of the prompt tone after the communication between the MG and the MGC is interrupted, use this parameter.	0: indicates the busy tone.
23	Indicates the length of the timer for synchronizing the	35s

Parameter	Description	Default Setting
	port status. To configure the length of the timer for synchronizing the port status, use this parameter.	
24	Indicates the maximum value of the Real-Time Transport Protocol (RTP) termination ID.	-
25	Indicates the maximum random value for the protection against avalanche of the H.248 interface.	-
26	Indicates the type of local blocking play tone.	0: indicates the busy tone.
27	Indicates the type of remote blocking play tone.	0: indicates the busy tone.
28	Indicates the duration of the howler tone.	60s
29	Indicates the duration of message waiting tone.	3s
30	Indicates the time limit of the alarm for extra long call.	60 minutes
31	Indicates whether to report the alarm for extra long call.	1: indicates that the alarm is not reported.
32	Min. auto registration interval of remotely-blocked port(s).	1800s
33	Whether MG heartbeat is shut down.	1: No, heartbeat is enabled

There are 17 software parameters (numbered from 0 to 16) of an MG interface that supports MGCP. Table 23-49 lists the configurable parameters, and the other parameters are reserved in the system.

Table 23-49 Software parameters of an MG interface that supports MGCP

Parameter	Description	Default Setting
1	Indicates whether the ongoing call is maintained if the communication between the MG interface and the MGC is abnormal. To maintain the ongoing call when the communication	1: disconnects all the calls at once.

Parameter	Description	Default Setting
	between the MG interface and the MGC is abnormal, use this parameter.	
2	Indicates whether the MG interface supports dual homing. To configure whether an MG interface supports dual homing, use this parameter. If the MG interface does not support dual homing (value: 1), even if the secondary MGC is configured, the MG interface does not switch to register with the secondary MGC when the MG interface fails to register with the primary MGC.	0: indicates that dual homing is supported.
3	Indicates whether the heartbeat message between the MG and the MGC is disabled. To configure whether the heartbeat message between the MG and the MGC is disabled, use this parameter.	1: indicates that the heartbeat message is not disabled.
4	Indicates whether a wildcard is used for the registration of the MG interface. To configure whether a wildcard is used for the registration of an MG interface, use this parameter. Using a wildcard for registration can reduce the quantity of messages between the MG interface and the MGC. When a wildcard is not used for registration, all the terminals on the MG interface need to register with the MGC in order. The registration without a wildcard is also called "single-endpoint registration".	0: indicates that a wildcard is used.
5	Indicates the MGC type. To select the MGC of a different type, use this parameter.	0

Parameter	Description	Default Setting
6	Indicates the maximum time threshold for responding to the heartbeat messages. To configure the maximum times for transmitting the heartbeat message continuously, use this parameter.	3
7	Indicates whether to report the heartbeat with the MG as an endpoint. To configure whether to report the heartbeat with the MG as an endpoint, use this parameter.	0: indicates that reporting the heartbeat with the MG as an endpoint is not supported.
10	Indicates the point-to-point (P2P) fault reporting. To configure whether the MG interface supports the P2P fault reporting from the ISDN terminals, use this parameter.	0: indicates that the P2P fault is reported.
11	Indicates the point-to-multipoint (P2MP) fault reporting. To configure whether the MG interface supports the P2MP fault reporting from the ISDN terminals, use this parameter.	1: indicates that the P2MP fault is not reported.
12	Indicates the type of local blocking play tone.	0: indicates the busy tone.
13	Indicates the type of remote blocking play tone.	0: indicates the busy tone.
14	Indicates the RTP filtering switch of the MGCP interface. To configure whether the RTP filtering function is enabled, use this parameter. When the RTP filtering function is enabled, only the media stream from the specific ports can be received.	1: indicates that the RTP filtering function is not enabled.
15	Indicates the duration of the howler tone.	60s
16	Whether the timer symbol "T" follows the number string	0: Yes

Parameter	Description	Default Setting
	reported by the signaling.	

Procedure

- When the system protocol is H.248, perform the following operations to configure the software parameters of an MG interface.
  a. In the global config mode, run the **interface h248** command to enter the MG interface mode.
  b. Run the **mg-software parameter** command to configure the software parameters listed in the data plan.
  c. Run the **display mg-software parameter** command to check whether the software parameters of the MG interface are the same as those in the data plan.
- When the system protocol is MGCP, perform the following operations to configure the software parameters of an MG interface.
  a. In the global config mode, run the **interface mgcp** command to enter the MG interface mode.
  b. Run the **mg-software parameter** command to configure the software parameters listed in the data plan.
  c. Run the **display mg-software parameter** command to check whether the software parameters of the MG interface are the same as those in the data plan.

----End

Example

To configure software parameter 11 of H.248-based MG interface 0 to 1 so that the MG interface supports emergency standalone and allows only internal calls, do as follows:

```
huawei(config)#interface h248 0
huawei(config-if-h248-0)#mg-software parameter 11 1
huawei(config-if-h248-0)#display mg-software parameter 11
  ------------------------------------------------
  Interface Id:0         para index:11  value:1
  ------------------------------------------------
 APPENDIX:
  ------------------------------------------------
   Interface software parameter name:
   11: Stand alone support
      0: None
      1: Inner
      2: Emergency
      3: Both
```

(Optional) Configuring the TID Format of an MG Interface

The TID format determines how various user terminal IDs on the MG interface are generated.

Prerequisites

NOTICE

The configuration of the TID format is relatively complicated. The information such as the syntax rules with which the terminal prefix must comply and the requirements for the character string of the TID template is defined in the protocol, and is not described here. This topic describes only some basic information. It is recommended that you refer to the TID description in ITU-T H248.1 (applicable to H.248) or RFC 3435 (applicable to MGCP) before configuring the TID format.

Context

The TID format consists of the terminal prefix and the TID template. The TID template defines the generation mode of the TID excluding the terminal prefix. A TID consists of a terminal prefix and a character string generated by a TID template.

The TID templates that are bound to various types of users on the MG interface determine whether the users support terminal layering.

- If the parameter list of the TID template includes only keyword "G", the TID template is used by the non-layering users. Users bound with this template do not support terminal layering.
- If the parameter list of the TID template includes only keywords "F", "S", "P", "B" ("B" is not available to PSTN users), the TID template is used by the layering users. Users bound with this template support terminal layering.

NOTE

The meaning of each keyword is as follows:

- F indicates the subrack ID.
- S indicates the slot ID.
- P indicates the port ID.
- B indicates the B channel ID (only for ISDN BRA and ISDN PRA terminals).
- G indicates the general permanent termination index.
- R indicates the RTP ephemeral termination index (only for the RTP ephemeral termination, which exists only when the system protocol is H.248. This termination is not involved unless special remarks are provided.)

When adding a user that supports terminal layering, you cannot and do not have to specify the parameter **terminalid** because the system automatically allocates a terminal ID. When adding a user that does not support terminal layering, you must specify the parameter **terminalid**.

You can run the **display tid-format(h248)** command or **display tid-format(mgcp)** command to query the TID formats of various types of users on the current MG interface. In the query result, **template-index** indicates the index of the TID template that is bound to the type of users. Then, run the **display tid-template** command to check whether the TID template supports the layering configuration. Hence, you can check whether the user supports terminal layering.

Precaution

- There are 19 default TID templates (templates 0-18) in the system. The default TID templates can be referenced, but cannot be added, modified, or deleted.
- The configuration of terminal layering on the MG must be the same as that on the MGC.
- If certain type of terminals exists on an interface and the interface is not disabled, the terminal prefix of this type of terminals cannot be modified.
- If a certain type of terminals exists on an interface, the TID format (including the terminal prefix and the index of the TID template) of this type of terminals cannot be changed.
- The terminal prefix must comply with the following syntax rules: The prefix must not exceed 64 characters. Only letters, digits, "_", and "/" are the characters allowed for input. The first character must be a letter, and the last character must not be a digit.
- The length of the TID, which is generated by combining the TID template and the terminal prefix that you configured, must not exceed 64 characters.

Procedure

- If the system protocol is H.248, to configure the TID format on the current MG interface, do as follows:
  a. Run the **display tid-template** command to query the information about the default TID template of the system.
  b. If the default TID template cannot meet the service requirements, run the **tid-template add** command to add a TID template that meets the service requirements. If the default TID template meets the service requirements, proceed to c.
  c. Run the **interface h248** command to enter the H.248 mode.
  d. Configure the TID templates and the terminal prefix of various types of users on the current MG (VAG).

NOTICE

The terminal prefix of PSTN, BRA and PRA must be all identical or unique.

    - In the H.248 mode, run the **tid-format rtp** command to configure the TID template and the terminal prefix of the RTP ephemeral termination.
    - In the H.248 mode, run the **tid-format pstn** command to configure the TID template and the terminal prefix of the PSTN user.
    - In the H.248 mode, run the **tid-format bra** command to configure the TID template and the terminal prefix of the ISDN BRA user.
    - In the H.248 mode, run the **tid-format pra** command to configure the TID template and the terminal prefix of the ISDN PRA user.

  e. Run the **display tid-format(h248)** command to check whether the TID templates and the terminal prefixes of various types of users on the current MG interface are the same as those in the data plan.
- If the system protocol is MGCP, to configure the TID format on the current MG interface, do as follows:

a. Run the **display tid-template** command to query the information about the default TID template of the system.
b. If the default TID template cannot meet the service requirements, run the **tid-template add** command to add a TID template that meets the service requirements. If the default TID template meets the service requirements, proceed to c.
c. Run the **interface mgcp** command to enter the MGCP mode.
d. Configure the TID templates and the terminal prefix of various types of users on the current MG (VAG).
   - In the MGCP mode, run the **tid-format pstn** command to configure the TID template and the terminal prefix of the PSTN user.
   - In the MGCP mode, run the **tid-format bra** command to configure the TID template and the terminal prefix of the ISDN BRA user.
   - In the MGCP mode, run the **tid-format pra** command to configure the TID template and the terminal prefix of the ISDN PRA user.
e. Run the **display tid-format(mgcp)** command to check whether the TID templates and the terminal prefixes of various types of users on the current MG interface are the same as those in the data plan.

----End

Example

Assume that in the H.248 mode, the terminal prefix of the PSTN user on MG interface 1 is aln/, and the layering TID template 3 is used. To add a PSTN user on port 0/2/0 and check whether the system automatically allocates a TID generated according to the template, do as follows:

```
huawei(config)#display tid-template 3//Query the information about TID template 3
  -------------------------------------------------
  Index     : 3
  Format    : %u/%u/%u
  Para-list  : F+1,S+1,P+1  //The parameter list of the TID template includes keyword
"F", "S",
                //and "P", which indicates that this template supports terminal layering.
  Name      : Aln_Not_Fixed_1
  -------------------------------------------------

huawei(config)#interface h248 1
huawei(config-if-h248-1)#tid-format pstn prefix aln template 3
huawei(config-if-h248-1)#quit
huawei(config)#esl user
huawei(config-esl-user)#mgpstnuser add 0/2/0 1
huawei(config-esl-user)#display mgpstnuser 0/2/0
{ <cr>|endframeid/slotid/portid<S><1,15> }:

  Command:
         display mgpstnuser 0/15/0
  ------------------------------------------------------------------------
  F  /S /P    MGID    TelNo          Priority PotsLineType TerminalID
  ------------------------------------------------------------------------
  0  /2 /0    1       -              Cat3     DEL         aln/1/3/1
           //The system allocates the terminal ID according to the TID format.
```

```
------------------------------------------------------------------------
```

Enabling an MG Interface

Enabling an MG interface is to reset an MG interface to make the MG interface register with the MGC (or to make the modified attributes of the MG interface take effect) after the configuration of the MG interface is complete, so that the MG interface can work in the normal state.

Precaution

CAUTION

For the MG interface that has been in service, this operation interrupts the ongoing services carried on the MG interface. Hence, exercise caution when performing this operation.

Procedure

- Enable the MG interface that adopts the H.248 protocol.
  a. Run the **interface h248** command to enter the H.248 mode.
  b. Run the **reset coldstart** command to enable the MG interface.
  c. Run the **quit** command to return to the global config mode, and then run the **display if-h248 all** command to check whether the MG interface is in the normal state.
- Enable the MG interface that adopts the MGCP protocol.
  a. Run the **interface mgcp** command to enter the MGCP mode.
  b. Run the **reset** command to enable the MG interface.
  c. Run the **quit** command to return to the global config mode, and then run the **display if-mgcp all** command to check whether the MG interface is in the normal state.

----End

Example

To enable H.248-based MG interface 0, do as follows:

```
huawei(config)#interface h248 0
huawei(config-if-h248-0)#reset coldstart
  Are you sure to reset MG interface?(y/n)[n]:y
huawei(config-if-h248-0)#quit
huawei(config)#display if-h248 all
------------------------------------------------------------------------------
MGID    TransMode State        MGPort MGIP         MGCPort MGCIP/DomainName
------------------------------------------------------------------------------
0       UDP      Normal       2944   10.10.10.11    2944   10.10.20.11
------------------------------------------------------------------------------
```

To enable MGCP-based MG interface 0, do as follows:

```
huawei(config)#interface mgcp 0
huawei(config-if-mgcp-0)#reset
  Are you sure to reset MG interface?(y/n)[n]:y
huawei(config-if-mgcp-0)#quit
huawei(config)#display if-mgcp all
  ------------------------------------------------------------------------------
  MGID    State         MGPort MGIP           MGCPort MGCIP/DomainName
  ------------------------------------------------------------------------------
  0       Normal        2727   10.10.10.11    2727    10.10.20.11
  1       Wait ack      2527   10.10.10.12    2727    10.10.20.12
  ------------------------------------------------------------------------------
```

23.22.2 Configuring the IUA Link

This topic describes how to configure the IUA link for signaling transmission between the Access node and MGC in the VoIP ISDN service.

Context

- Simple Control Transmission Protocol (SCTP) is a connection-oriented protocol. Its most fundamental function is to provide reliable transmission for interaction messages between the Access node and MGC. The SCTP protocol implements services based on the association between two SCTP endpoints. SCTP can be regarded as a transmission layer. Its upper layer is called SCTP subscriber, and its lower layer is the IP network.
- The IUA link is the carrier of the interaction signaling between the Access node and MGC.

Adding an IUA Link Set

This topic describes how to add an IUA link set. When configuring the VoIP ISDN service, you need to configure the IUA link to carry the Q.931 call signaling. Before adding an IUA link, you must add a corresponding link set. Otherwise, the link cannot be added.

Context

- The system supports the configuration of a maximum of IUA link sets.
- After a link set is configured successfully, it is in the deactivated state by default.

Procedure

In global config mode, run the command**sigtran** to enter the Sigtran mode.

Step 1 Run the **iua-linkset add** command to add an IUA link set.

Step 2 You can run the **display iua-linkset attribute** command to check whether the configured IUA link set information is the same as the data plan.

----End

Example

Assume that the link set ID is 0, working mode of the link set is active/standby mode, pending duration is 20s, prefix of the IID is b/, IID generation mode is using the binary value that is automatically generated in ffsspp mode, namely, parameter 2. To add such an IUA link set, do as follows:

```
huawei(config)#sigtran
```

Adding an IUA Link

This topic describes how to add an IUA link. After the link set is added, you can add an IUA link to carry the Q.931 call signaling for the ISDN user.

Prerequisites

The IUA link set must be added.

Context

- Make sure that a minimum of one item in the local port ID, local IP address, remote port ID, and remote IP address of a link is different from the corresponding item of other links.
- Only two links can be configured in the same link set. In addition, the local IP addresses of the two links must be the same.

Procedure

(This step is not required if the command line interface is already in the Sigtran mode.) In global config mode, run the **sigtran** command to enter the Sigtran mode.

Step 1 Run the **iua-link add** command to add an IUA link.

Step 2 You can run the **display iua-link attribute** command to check whether the configured IUA link information is the same as the data plan.

----End

Example

Assume that the link ID is 0, link set ID is 0, local port ID is 1402, local IP address is 10.10.10.10, remote port ID is 1404, and remote IP address 1 is 10.10.10.20. To add such a link, do as follows:

```
huawei(config)#sigtran
huawei(config-sigtran)#iua-link add 0 0 1402 10.10.10.10 1404 10.10.10.20
huawei(config-sigtran)#display iua-link attribute
{ <cr>|linkno<L> }:

 Command:
       display iua-link attribute
 ------------------------------------------------------------------------
 LinkNo             :  0
 LinksetNo          :  0
 Local port         :  1402
```

```
Local IP address     :  10.10.10.10
Remote port          :  1404
Remote IP address    :  10.10.10.20
Remote IP address 2  :  -
Priority             :  0
--------------------------------------------------------------------------
```

23.22.3 Configuring the VoIP ISDN User

This topic describes how to configure the VoIP ISDN user. After the MG interface is configured, you can add the VoIP ISDN user on this interface to implement the VoIP ISDN service.

Configuring the Attributes of the E1 Port

This topic describes how to configure the attributes of the E1 port to ensure that the ISDN PRA port meets the actual application requirements.

Context

You can configure the impedance, line coding mode, and working mode of the E1 port.

Default Configuration

Table 23-50 lists the default values of the E1 port. When configuring the attributes of the E1 port, you need to modify the values according to the service requirements.

Table 23-50 Default values of the E1 port

Parameter	Default Setting
Port impedance	E1 mode: 75 ohm
Line coding mode	E1 mode: HDB3
CRC4	Enable
The mode for digital section access	Digital
Signaling type	CCS

Procedure

In global config mode, run the **interface edt** command to enter the EDT mode.

Step 1 (Optional) Run the **e1port impedance** command to configure the impedance of an E1 port.

Step 2 (Optional; perform this step when you need to modify the line coding mode of the port) Run the **e1port line-code** command to configure the line coding mode of the E1 port.

NOTE
In E1 mode, the system supports two line coding modes, namely, HDB3 and AMI.

Step 3 (Optional) Run the **e1port crc4** command to configure the CRC4 function of an E1 port.

Step 4 (Optional) Run the **e1port attribute set** command to configure the digital section access mode of an E1 port.

Step 5 (Optional) Run the **e1port signal** command to configure the signaling type of an E1 port.

----End

Example

Assume that the E1 ports on ISDN PRA board work in HDB3 line encoding mode, the CRC4 function is enable, To configure such E1 ports, do as follows:

```
huawei(config)#interface edt 0/1
huawei(config-if-edt-0/1)#e1port line-code 1 HDB3
huawei(config-if-edt-0/1)#display e1port line-code 1
  ---------------------
  F/S/P   linecode
  ---------------------
  0/1/1   HDB3
  ---------------------
huawei(config-if-edt-0/1)#e1port crc4 1 enable
huawei(config-if-edt-0/1)#display e1port attribute 1
  ------------------------------------------------------------
   F/S/P    Signaltype    CRC4    Impedance    Accessmode
  ------------------------------------------------------------
   0/1/1    CCS          Enable  75          Digital
  ------------------------------------------------------------
```

Configuring the ISDN PRA User Data

This topic describes how to configure the ISDN PRA user data on the H.248 interface (the data must be the same as the corresponding data on the MGC) so that the ISDN PRA user can access the network to use the ISDN PRA service.

Prerequisites

The ISDN PRA service board must be inserted into the planned slot correctly and the board must be in the normal state.

Default Configuration

Table 23-51 lists the default settings of the attributes of the ISDN PRA user. When configuring these attributes, you can modify the values according to the service requirements.

Table 23-51 Default settings of the attributes of the ISDN PRA user

Parameter	Default Settings
Priority of the ISDN PRA user	cat3: (common user)
Flag of reporting the UNI fault of the ISDN PRA user	Disable
Sub-channel active mask of the ISDN PRA	255.255.255.255

Parameter	Default Settings
user	
Threshold for the number of auto recoveries from deterioration faults	20

Procedure

In global config mode, run the **esl user** command to enter the ESL user mode.

Step 1 Run the **mgprauser add** command to add an ISDN PRA.

- When **iid-map** in the **iua-linkset add** command is configured to 1, interfaceid must be configured and be different from the interfaceid of other users in the same link set.
- The terminal ID of an ISDN PRA user must be unique.
- Each ISDN PRA user occupies 32 terminal IDs. You need to input only the first terminal ID when adding an ISDN PRA user. If the MG interface does not support the terminal layering function, plus (or minus) at least 32 to (or from) the terminal ID of the previous ISDN PRA user when adding an ISDN PRA user, and use the result as the first terminal ID of the current ISDN PRA user.
- If the MG interface supports the terminal layering function, the terminal ID cannot be configured when an ISDN PRA user is added on the MG interface. The system automatically allocates a terminal ID for the user.

Step 2 Run the **display mgprauser** command to check whether the ISDN PSTN user data is the same as the data plan.

Step 3 (Perform this step only when you need to modify the attributes of the ISDN PRA user.) Run the **mgprauser attribute set** command to configure the attributes of the ISDN PRA user.

Step 4 (Perform this step only after you modify the attributes of the ISDN PRA user.) Run the **display mgprauser** command to query whether the configured attributes of the ISDN PRA user are the same as the data plan.

----End

Example

Assume that the MG interface ID is 0, link set ID is 0, IUA port ID is 0 (iid-map is 1), terminal ID is 223 (does not support the terminal layering function), user priority is cat2, UNI alarm report flag is enable, sub-channel active mask is 255.255.255.255, and the threshold for the number of auto recoveries from L1 is 30. To add such an ISDN PRA user connected to the 0/1/0 port, do as follows:

```
huawei(config)#esl-user
huawei(config-esl-user)#mgprauser add 0/1/0 0 0 interfaceid 0 terminalid 223
huawei(config-esl-user)#display mgprauser 0/1
  ------------------------------------------------------------
  F  /S /P /B   MGID    LinkSetNo   InterfaceID   TerminalID
  ------------------------------------------------------------
  0/1/0 /0   0       0          0           A223
  ------------------------------------------------------------
```

```
huawei(config-esl-user)#mgprauser attribute set 0/1/0 priority cat2 unireport en
able activemask 255.255.255.255 auto-resume-limit 30
huawei(config-esl-user)#display mgprauser attribute 0/1/0
{ <cr>|endframeid/slotid/portid<S><Length 1-15> }:

  Command:
         display mgprauser attribute 0/1/0
  ------------------------------------------------------------------------------
  F  /S /P                : 0/1/0
  UNIreport                : enable
  Prior                   : Cat2
  Mask of sub channel       : 255.255.255.255
  Auto reservice times/limit : 0/30
   ------------------------------------------------------------------------------
```

(Optional) Configuring the System Parameters

This topic describes how to configure the system parameters including the overseas version flag and message waiting indication (MWI) mode according to the local standards to ensure that the response of the user terminal complies with the local standards.

Procedure

Run the **system parameters** command to configure the system parameters.

Step 1 Run the **display system parameters** command to check whether the system parameters are the same as those in the data plan.

----End

Example

To configure the overseas version flag (system parameter 1) to Hong Kong (parameter value 1), do as follows:

```
huawei(config)#system parameters 1 1
huawei(config)#display system parameters 1
  ------------------------------------------------------------------------------
  Parameter name index: 1    Parameter value: 1
  Mean: Overseas version flag, 0:China, 1:HongKong, 2:Brazil, 3:Egypt, 4:
Singapore, 5:Thailand, 6:France, 7:Britain MSFUK, 8:Britain ETSI, 9:Bulgaria,
10:Reserved, 11:Austria, 12:Hungary, 13:Poland
  ------------------------------------------------------------------------------
```

(Optional) Configuring the Overseas Parameters

By default, the overseas parameters are configured according to the Chinese standards. In the actual service configuration, the attributes such as the upper and lower thresholds of the flash-hooking duration can be configured according to the local standards to ensure that the response of the user terminal complies with the local standards.

Procedure

Run the **oversea parameters** command to configure the overseas parameters.

Step 1 Run the **display oversea parameters** command to check whether the overseas parameters are the same as those in the data plan.

----End

Example

To set the upper flash-hooking threshold (overseas feature parameter 0) to 800 ms (in compliance with the Hong Kong standard) and the lower flash-hooking threshold (overseas feature parameter 1) to 100 ms (in compliance with the Hong Kong standard), do as follows:

```
huawei(config)#oversea parameters 0 800
huawei(config)#oversea parameters 1 100
huawei(config)#display oversea parameters
{ <cr>|name<U><0,99> }:

 Command:
        display oversea parameters
 ------------------------------------------------------------------------------
 Parameter name index: 0    Parameter value: 800
 Mean: Hooking upper threshold(ms), reference: China:350, HongKong:800
 ------------------------------------------------------------------------------
 Parameter name index: 1    Parameter value: 100
 Mean: Hooking lower threshold(ms), reference: China:100, HongKong:100
 ------------------------------------------------------------------------------
 Parameter name index: 2    Parameter value: 0
 Mean: Flag of applying PARKED LINE FEED or not when user port is locked, 0:not
apply, 1:apply
 ------------------------------------------------------------------------------
 Parameter name index: 3    Parameter value: 0
 Mean: The detect time of flash upper limit to onhook, default value: 0ms
 ------------------------------------------------------------------------------
 Parameter name index: 4    Parameter value: 0
 Mean: DTMF Detector Tuning, 0:Q.24, 1:Russia
------------------------------------------------------------------------------
 Parameter name index: 5    Parameter value: 0
 Mean: Flag of changing TEI value for software switch, 0:no change, 1:set TEI
value to 1, 2:set TEI value to 0. Default: 0
 ------------------------------------------------------------------------------
```

23.23 Configuring the R2 Service

With the R2 access technology, the Access node provides access services on common twisted pair cables when interconnecting with the PBX using R2 signaling.

Prerequisites

- The operation mode of the EDTB board must be configured as service mode, which can be configured using the **board workmode** command.

- The working mode of the EDTB board must be configured as voice mode, which can be configured using the **runmode** command.
- The signaling type of ports on the EDTB board must be configured as channel associated mode, which can be configured using the **e1port signal** command.

For the H.248-based MoIP service:

- An MG interface has been configured. For details, see 23.20.1 Configuring an MG Interface.

For the SIP-based MoIP service:

- A SIP interface has been configured. For details, see 23.17.1 Configuring an SIP Interface.

Context

- R2 signaling is channel associated signaling (CAS), which is international standard signaling based on E1 digital network.
- The Access node connects the PBX and NGN network using R2 signaling, achieving transition from the PSTN network to the NGN network.

Procedure

Run the **r2 profile** command to add an R2 profile. Define the R2 signaling with a specific feature as an R2 profile which can be used when an R2 user is added.

Step 1 (Optional) Run the **profile attribute** command to configure the signaling type of the R2 profile.

Step 2 (Optional) Configure adaptation data of the R2 profile.

The ITU-T Q.400-Q.490 standard has defined R2 signaling standard, but different countries and regions implement R2 signaling in different ways. You do not need to change parameter values if the parameter values defined in the signaling standard of a country are consistent with the default values defined by the Access node. Otherwise, you need to change the parameter values based on actual conditions.

- Run the **address-receive attribute** command to configure the receive attribute of R2 addresses.
- Run the **address-send attribute** command to configure the transmit attribute of R2 addresses.
- Run the **profile attribute** command to configure the signaling type of the R2 profile.
- Run the **line-signaling attribute** command to configure the R2 line signaling attribute.
- Run the **register-signaling attribute** command to configure the R2 register signaling attribute.

Step 3 (Optional) Run the **multi-r2-adapt add** command to add parameters of the state machine of register signaling and line signaling in adaptation profiles for multiple countries. To comply with the R2 standards of different countries, parameter configuration for the R2 state machine is added, so that mappings between logical commands and physical commands can be changed by changing configurations without adding logical commands.

Step 4 Run the **mgr2user add** command (for H.248 protocol) or **sipr2user add** command (for SIP protocol) to add an R2 user.

Step 5 (Optional) Run the **mgr2user attribute set** command (for H.248 protocol) or **sipr2user attribute set** command (for SIP protocol) to set the priority of R2 users. When congestion occurs, packets of the user with a high priority is forwarded first.

----End

Example

For example, configure R2 users at the 0/2/1 port when using SIP protocol. Parameters are shown in Table 23-52.

Table 23-52 Data plan for R2 user configuration

Configuration Item	Data
R2 profile	0
Signaling type of the R2 profile	10
Wait-answer-time	200s
Wait-protect-time	300 ms
SIP interface	0
Subrack ID/slot ID/port number	0/2/1
The terminal priority of an R2 user	cat2

```
huawei(config)#r2 profile 0
  Are you sure to add r2 profile?(y/n)[n]:y
huawei(config-r2-0)#profile attribute name normal signaling-type 10
huawei(config-r2-0)#line-signaling attribute wait-answer-time 200 wait-protect-time
300
huawei(config-r2-0)#quit
huawei(config)#esl user
huawei(config-esl-user)#sipr2user add 0/2/1 0 0
huawei(config-esl-user)#sipr2user attribute set 0/2/1 priority cat2
```

23.24 Configuring the H.248/MGCP-based FoIP Service

This topic describes how to configure the H.248/MGCP-based FoIP service.

Prerequisites

The VoIP service must be configured. For details, see 23.20 Configuring the VoIP PSTN Service (H.248-based or MGCP-based).

The voice service port used for the FoIP service must be in the normal state, and the voice communication on the port must be normal.

Context

Fax over Internet Protocol (FoIP) provides fax services on IP networks or between IP networks and PSTN networks. The FoIP service can be classified from two aspects:

- In terms of coding mode, the fax mode can be transparent fax (G.711 coding) or T.38 fax (T.38 coding).
- In terms of the participation of the MGC, one is the softswitch controlled flow, and the other is the self-switch flow (controlled by the gateway itself).

Data preparation

Before the data configuration, it is recommended that you plan the working mode of the FoIP service on the entire NGN network to ensure that the configurations on the entire network are consistent.

Table 23-53 Fax flows

Item	Flow	Remarks
Coding negotiation mode	Negotiation	The gateway negotiates the coding mode with the MGC through signaling.
	Self-switch	The gateway determines the coding mode to be adopted.
Coding mode	Transparent transmission fax	The G.711 coding mode is adopted.
	T.38 flow	The T.38 coding mode is adopted.
Negotiation flow	V2 flow (auto-negotiation flow)	The V2 flow is adopted as the fax/modem flow.
	V3 flow	The V3 flow is adopted as the fax/modem flow.
	V5 flow	The V5 flow is adopted as the fax/modem flow.

NOTE
V2, V3, and V5 flows refer to the versions of the fax/modem flow, which is defined by Huawei. If self-switch is adopted as the coding negotiation mode, the negotiation flow does not need to be configured.

Default Configuration

Table 23-54 lists the default settings of the FoIP flow.

Table 23-54 Default settings of the FoIP flow

Item	Default Setting
Coding negotiation mode	Negotiation
Coding mode	Transparent transmission fax

Item	Default Setting
Negotiation flow	V3 flow
Enable packet interval of fax and modem to use only 10 ms or not	Disable
RFC2198 startup mode	DisableRfc2198SmartStartup
Event transmit mode	ControlledByMGC

Procedure

Configure public fax and modem parameters.

In the global config mode, run the **fax-modem parameters negomode** command to configure public fax and modem parameters. The purpose of this step is to configure the coding negotiation mode. Two options are available: negotiation and self-switch.

Step 1 Configure the fax coding mode and negotiation flow.

In the global config mode, run the **fax parameters** command to configure the fax transmission mode. There are three key parameters, which are described as follows:

- **transmode**: The value 0 indicates the transparent transmission mode with the G.711 coding; the value 1 indicates the T.38 mode, which is a coding mode dedicated to the fax service. The default value is 0.
- **flow**: Options are V2, V3, and V5. The default value is V3.
- **is-port+2**: This parameter should be consistent with the T.38 fax port configured on the peer MGC. When **transmode** is T.38 and **flow** is V2, this parameter must be configured.

NOTICE

In the high-speed fax mode, the fax mode cannot be configured as the auto-negotiation (V2 flow) T.38 mode or the self-switch transparent transmission mode. If the fax mode is the auto-negotiation (V2 flow) T.38 mode or the self-switch transparent transmission mode, modify the configuration according to step 1 and step 2.

Step 2 Query the common parameters of the fax and modem or the fax parameter configuration.

1. Run the **display fax-modem parameters** command to query the fax negotiation flow.
2. Run the **display fax parameters** command to query the fax coding mode.

----End

Example

To configure the negotiation mode of the FoIP service on the MA5600T/MA5603T/MA5608T to negotiation, enable 10 ms packetization, enable RFC2198 smart startup mode, configure the event transfer mode of fax to RFC2833, and the fax working mode of the

MA5600T/MA5603T/MA5608T to thoroughly, and configure the fax flow to V2 flow, do as follows:

```
huawei(config)#fax-modem parameters negomode selfswitch packet-interval-10ms en
able rfc2198-start-mode enableRfc2198SmartStartup transevent rfc2833
huawei(config)#fax parameters flow v2 workmode thoroughly
huawei(config)#display fax-modem parameters
  ---------------------------------------------------------------------------
  Negomode             : Self switch
  Packet-interval-10ms   : Enable
  Rfc2198-start-mode    : Enable Rfc2198SmartStartup
  TransEvent           : RFC2833
  Vbd-codec            : G.711A
  Vbd-payload-type      : Static
  ---------------------------------------------------------------------------

huawei(config)#display  fax parameters
  ---------------------------------------------------------------------------
  FAX transfers mode               :Thoroughly
  T38 Fax Port                    :RTP port
  FAX flow                        :V2 Flow
  ---------------------------------------------------------------------------
```

23.25 Configuring the SIP-based FoIP Service

This topic describes how to configure the SIP-based FoIP service.

Prerequisites

- The SIP-based VoIP service must be configured. For details, see 23.17 Configuring the VoIP PSTN Service (SIP-based).
- The voice service port used for the FoIP service must be in the normal state, and the voice communication on the port must be normal.

Context

According to the fax coding mode, the FoIP service is classified into two modes:

- Transparent transmission fax: uses the G.711 coding
- T.38 fax: uses the T.38 coding

In the fax service application, according to whether the SIP signaling is involved in controlling the transmission, the FoIP service is classified into two modes:

- Negotiate mode, in which the SIP signaling is involved in controlling the transmission
- Self-switch mode, in which the SIP signaling is not involved in controlling the transmission

Data preparation

Before the data configuration, it is recommended that you plan the working mode of the FoIP service on the entire IMS network to ensure that the configurations on the entire network are consistent.

Table 23-55 Fax flows

Item	Flow	Remarks
Coding negotiation mode	Negotiate	The gateway negotiates the coding mode with the IMS through SIP signaling.
	Self-switch	The gateway determines the coding mode to be adopted.
Coding mode	Transparent transmission fax	The G.711 coding mode is adopted.
	T.38 flow	The T.38 coding mode is adopted.

Default Configuration

Table 23-56 lists the default settings of the FoIP flow.

Table 23-56 Default settings of the FoIP flow

Item	Default Setting
Coding negotiation mode	negotiate
Coding mode	Transparent transmission fax

Procedure

Configure the common parameters of fax and modem.

1. In the global config mode, run the **interface sip** command to enter the SIP interface mode.
2. In the SIP interface mode, run the **fax-modem parameters negomode** command to configure the coding negotiation mode.

   The purpose of this step is to configure the coding negotiation mode. Key parameter **negomode**: includes the self-switch mode and the negotiate mode. By default, the negotiate mode is adopted.:

Step 1 Configure the fax coding mode.

In the SIP interface mode, run the **fax parameters** command to configure the fax transmission mode. There is only one key parameter, which is described as follows:

transmode: The value 0 indicates the transparent transmission mode with the G.711 coding; the value 1 indicates the T.38 mode, which is a coding mode dedicated to the fax service. By default, the value 0 is adopted.

Step 2 Query the common parameters of the fax and modem or the fax parameter configuration.

1. Run the **display fax-modem parameters** command to query the fax negotiation flow.
2. Run the **display fax parameters** command to query the fax coding mode.

----End

Example

To configure the negotiation mode of the FoIP service on the SIP interface 0 to negotiate, enable 10 ms packetization, configure the RFC2198 negotiate mode to fixed start, RFC2198 start mode to smart2198 start, configure the event transfer mode of fax to fixed start, and the fax working mode to thoroughly, do as follows:

```
huawei(config-if-sip-0)#fax-modem parameters negomode negotiate packet-interval-
10ms enable rfc2198-negomode fixedstart rfc2198-startmode smart2198 transevent f
ixedsta
huawei(config-if-sip-0)#fax parameters transmode 0
huawei(config-if-sip-0)#display fax-modem parameters
 ---------------------------------------------------------------------------
 MGID               :0
 Nego-mode           :negotiate
 Packet-interval-10ms  :enable
 Rfc2198-nego-mode     :fixedstart
 Rfc2198-start-mode    :smart2198
 Vbd-codec            :G.711A
 Vbd-pt-type          :static
 Transfer-event        :fixedstart
 ---------------------------------------------------------------------------

huawei(config-if-sip-0)#display fax parameters
 --------------------------------------
 MGID     Transmode
 --------------------------------------
 0        Thoroughly
 --------------------------------------
```

23.26 Configuring the MoIP Service

This topic describes how to configure the H.248/MGCP/SIP-based MoIP service for transmitting the traditional narrowband modem data service over the IP network.

Prerequisites

For the H.248/MGCP-based MoIP service:

- The MG interface must be configured. For details, see 23.20.1 Configuring an MG Interface.
- The VoIP users must be configured. For details, see 23.20.2 Configuring the VoIP PSTN User.

For the SIP-based MoIP service:

- The SIP interface must be configured. For details, see 23.17.1 Configuring an SIP Interface.
- The VoIP users must be configured. For details, see 23.20.2 Configuring the VoIP PSTN User.

Context

The MoIP service can be transmitted in two modes:

- One is the transparent transmission mode, also called the voice-band data (VBD) transparent transmission. In this mode, the MG adopts the G.711 coding to encode and decode modem signals, and processes modem signals as common RTP data. In other words, the MG does not process modem signals, and the modem modulation signals are transparently transmitted over the IP network through the VoIP channel.
- The other is the redundancy mode, also called the relay mode.

NOTICE

Currently, the MA5600T/MA5603T/MA5608T supports the modem service only in the transparent transmission mode.

The modem event report mode is classified into the delay mode, direct mode, and high-speed signal immediate mode.

- In the delay mode, the MA5600T/MA5603T/MA5608T does not report the modem event immediately after receiving an event. Instead, it waits for a period of time until the event times out and no V21flag event is reported. In this manner, when the high-speed fax machines fail in the high-speed transmission (the modem mode on the host) negotiation, the low-speed transmission mode (the fax mode on the host) can still be used for transmitting data.
- In the direct mode, the MA5600T/MA5603T/MA5608T reports the modem event to the MGC immediately after receiving the event from the drive. To enable the MGC to quickly respond to a modem event, configure the modem event report mode to the direct mode.
- In the high-speed signal immediate mode, the MA5600T/MA5603T/MA5608T reports low-speed modem signals after a delay of 5.5s and reports high-speed modem signals without delay.

The configuration of the MoIP service is mainly the configuration of the modem event report mode and the transmission mode. The default settings are direct mode and transparent transmission mode. If configuration is required, you only need to configure the event report mode. This is because currently the MA5600T/MA5603T/MA5608T supports the modem service only in the transparent transmission mode. Hence, you do not need to configure the transmission mode.

Procedure

- Configure the H.248/MGCP-based modem event report mode.

  In the global config mode, run the **modem parameters eventmode** command to configure the modem event report mode. By default, the direct mode is used.
- Configure the SIP-based modem event report mode.

a. Run the **interface sip** command to enter the SIP interface mode.

b. (Optional) Run the **modem parameters transmode** command to configure the modem event report mode.

----End

Example

To enable the MA5600T/MA5603T/MA5608T to communicate with the MGC through H.248, and configure the transmission mode of the modem to transparent transmission and the modem event report mode to delay mode, do as follows:

```
huawei(config)#modem parameters eventmode 0
huawei(config)#save
```

23.27 Adding a POTS IP SPC

A semi-permanent connection (SPC) exclusively occupies a voice channel to meet the communication requirements and to ensure the communication quality for particular and vital access subscribers. To configure an IP SPC, configure the data such as the local IP address, local UDP port ID, remote IP address, and remote UDP port ID, and set up a direct IP connection between the two ends of the VoIP service. In this manner, the voice media data can be directly transmitted to the peer end.

Prerequisites

- The electrical switch is already switched to the VoIP daughter board.
- The IP address of the VLAN Layer 3 interface is already configured.
- The remote VLAN Layer 3 interface can be routed and reached from the local VLAN Layer 3 interface.

Procedure

Run the **voip** command to enter the VoIP mode.

Step 1 Run the **ip address** command to configure the media IP address for the VoIP service.

Step 2 Run the **quit** command to quit the VoIP mode.

Step 3 Run the **dsp-para-template add** command to configure the DSP parameter profile.

Step 4 Run the **spc** command to enter the SPC mode.

Step 5 Run the **ipspc add** command to add an SPC.

----End

Example

To add a POTS IP SPC with the parameters listed in Table 23-57, do as follows:

Table 23-57 Data plan for adding a POTS IP SPC

Item	Data
IP address of the local VLAN Layer 3 interface	192.168.0.10/24
Local UDP port ID	56988
IP address of the gateway	192.168.0.1
Media IP address	192.168.0.10
Remote IP address	192.168.1.100
Remote UDP port ID	56988
DSP parameter profile	• Profile name: **ecopen** • Status of echo suppression: 0 (enabled) • Jitter buffer mode: 1 (static mode) • Non-linear processing mode: 0 (disabled) • Status of silence compression: 1 (disabled) • DSP working mode: 0 (voice service)
Subrack ID/slot ID/port ID/channel ID	0/2/0/0 **NOTE** The channel ID for a PSTN subscriber must be 0.

```
huawei(config)#voip
huawei(config-voip)#ip address media 192.168.0.10 192.168.0.1
huawei(config-voip)#quit
huawei(config)#dsp-para-template add ecopen 0 1 0 1 0
huawei(config)#spc
huawei(config-spc)#ipspc add 0/2/0/0 local-ip 192.168.0.10 local-port 56988 remote-ip
 192.168.1.100 remote-port 56988 dsp-para-template ecopen
```

23.28 Configuring the IP Z Interface Extension Service

The following network typically applies to the scenario where Z interface extension private line service needs to be carried over the IP network for the headquarters (HQ) and branch offices of an enterprise after the PSTN network reconstruction. In the following configuration example, the FXO and FXS boards are added for the Z interface extension local MSAN and remote MSAN respectively, board attributes are configured, and IP semi-permanent connections (SPCs) of the IP Z interface extension type are created between the two boards, so that users connected to the FXS board are connected to the corresponding ports on the FXO board through the SPCs.

Prerequisites

IP Z interface extension is a technology proprietarily owned by Huawei. Therefore, the MSANs on the FXO side and the FXS side must be Huawei MSANs.

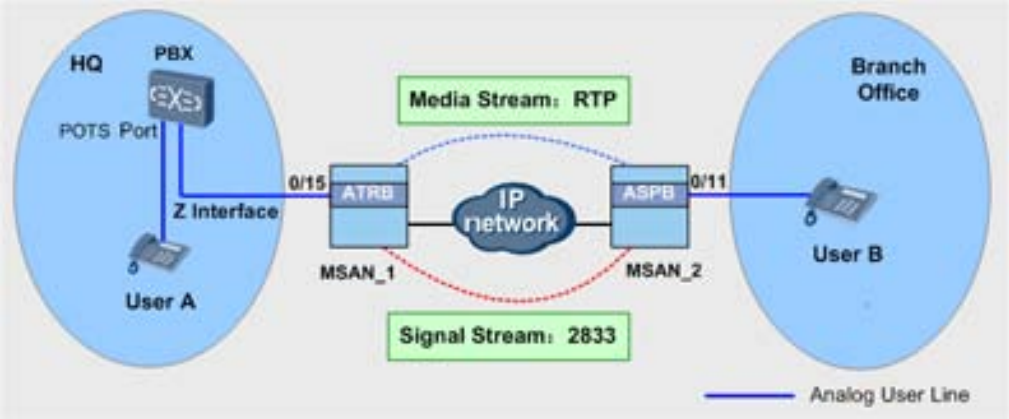
Application Scenario

NOTE

In this example, the MSANs at both ends of the IP network are Huawei MA5600T. The FXO board on MSAN_1 is the ATRB board, and the FXS board on MSAN_2 is the ASPB board. See more about the hardware support for the IP Z interface extension service , please refor to 23.13.1 Introduction to IP Z Interface Extension .

Figure 23-133 shows a sample network of the IP Z interface extension service. The ATRB board of MSAN_1 connects to user A through the local exchange (the PBX) by using the Z interface. User B connects to the ASPB board of MSAN_2. The two MSANs are connected through the IP network. SPCs of the IP Z interface extension type are created between corresponding ports of the ASPB and ATRB boards to implement the IP Z interface extension service.

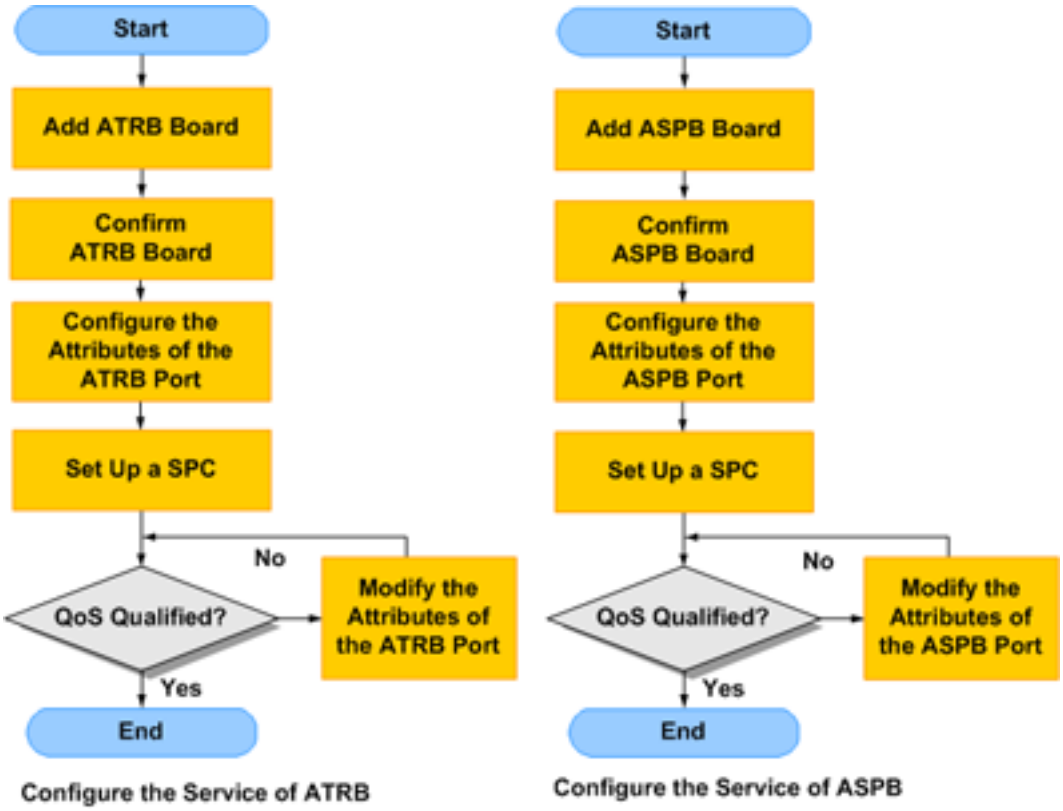
Figure 23-133 Network diagram of the IP Z interface extension service



Configuration Flow

Figure 23-134 shows the service configuration flows on the ATRB and ASPB boards.

Figure 23-134 Service configuration flows on the ATRB and ASPB boards



Procedure

Run the **board add** command to add an ATRB board on MSAN_1 and an ASPB board on MSAN_2.

Step 1 Run the **board confirm** command to confirm the ATRB board on MSAN_1 and the ASPB board on MSAN_2.

Step 2 on MSAN_1 and MSAN_2, run specially the **voip** command to enter the VOIP mode.

Step 3 on MSAN_1 and MSAN_2, run specially the **ip address** command to configure the IP address of the voice service.

Step 4 on MSAN_1 and MSAN_2, run specially the **quit** command to quit the VOIP mode.

Step 5 on MSAN_1, run the **fxoport** command to enter the FXO mode.

Step 6 on MSAN_1, run the **fxoport attribute set** command to configure port attributes for the ATRB board.

Step 7 on MSAN_1, run the**quit** command to quit the FXO mode.

Step 8 On MSAN_2, run the **pstnport** command to enter the PSTN mode.

Step 9 On MSAN_2, run the **pstnport_electric set** command to configure attributes of PSTN ports on the ASPB board.

Step 10 On MSAN_2, run the**quit** command to quit the PSTN mode.

Step 11 Run the **spc** command on MSAN_1 and MSAN_2 to enter the SPC configuration mode.

Step 12 Run the **ipspc add** command to configure SPCs of the IP Z interface extension type on MSAN_1 and MSAN_2.

----End

Example

Table 23-58 provides an example of related parameter configuration for the IP Z interface extension service.

Table 23-58 Data plan for the IP Z interface extension service

Configuration Item	Attribute	Data
ATRB board	Subrack ID/slot ID	0/15
ASPB board	Subrack ID/slot ID	0/11
VOIP address pools	Local IP address of media	10.10.10.2
	Remote IP address of media	10.10.10.3
	IP address of the gateway	10.10.10.1
ATRB port	Subrack ID/slot ID/port ID	0/15/0
	Shortest duration for ringing current detection	2
	Threshold for ringing current message detection	4
ASPB port	Subrack ID/slot ID/port ID	0/11/0
	Rx gain on a port	0
SPC	Local IP address of the IP SPC	10.10.10.2
	Local UDP port number of the IP SPC	57000
	Remote IP address of the IP SPC	10.10.10.3
	Remote UDP port number of the IP SPC	57004
	Number of channels occupied by the IP SPC	1

Configuration Item	Attribute	Data
	Signaling type of the IP SPC	2833 (Signaling is transmitted in the RFC2833 mode.)
	Initial ringing type of the IP SPC	Default value: 4 (This parameter is configured only on the port of the FXS board.)
	Cadence ringing type of the IP SPC	Default value: 0 (This parameter is configured only on the port of the FXS board.)

- Configuration procedure on MSAN_1

```
huawei(config)#board add 0/15 ATR
huawei(config)#board confirm 0/15
huawei(config)#voip
huawei(config-voip)#ip address media 10.10.10.2 10.10.10.1
huawei(config-voip)#quit
huawei(config)#fxoport
huawei(config-fxoport)#fxoport attribute set 0/15/0  ring-detect
min-ontime 2 max-offtime 4
huawei(config-fxoport)#quit
huawei(config)#spc
huawei(spc)#ipspc add 0/15/0/0 local-ip 10.10.10.2 local-port 57000
remote-ip 10.10.10.3 remote-port 57004 signal-type 2833
```

- Configuration procedure on MSAN_2

```
huawei(config)#board add 0/11 ASP
huawei(config)#board confirm 0/11
huawei(config)#voip
huawei(config-voip)#ip address media 10.10.10.3 10.10.10.1
huawei(config-voip)#quit
huawei(config)#pstnport
huawei(config-pstnport)#pstnport electric set 0/11/0 recvgain 0
huawei(config-pstnport)#quit
huawei(config)#spc
huawei(spc)#ipspc add 0/11/0/0 local-ip 10.10.10.3 local-port 57004
remote-ip 10.10.10.2 remote-port 57000 signal-type 2833 initialring 4
cadencering 0
```

NOTE

- You are advised to use default values for Tx gains of PSTN ports and FXO ports on the FXS board on MSAN_2 and those of PSTN ports on the PBX on MSAN_1. Do not change Tx gains of these ports to avoid the scenario in which dial tones cannot be cut because of Tx gain modification.
- You are advised to set the Rx gain of the PSTN port on the FXS board on MSAN_2 to its maximum 0 dB (default value is -7 dB), set that of the FXO port on the FXO board to -3 dB, and set that of the PSTN port on the PBX on MSAN_1 to -4 dB .
- In actual applications, use parameter **recvgain** in the **fxoport attribute set** command to configure the Rx gain of the FXO port and parameter **recvgain** in the **pstnport electric set** command to configure the Rx gain of the PSTN port on the FXS board.

- For details about the process and precautions of configuring ringing current detection parameters on the FXO port, see 23.13.5 Ringing and CLIP Services for IP Z Interface Extension Feature.
- After configuring the port attributes, run the **display fxoport attribute** command to query the configuring information of port attributes.
- After configuring the SPCs of the IP Z interface extension, run the **display ipspc** command to query the configuring information of IP SPCs.

After the configuration go into effect, user A and user B can communicate.

23.29 Configuring the Security and Reliability of the Voice Service

The security configuration of the voice service includes the H.248-based, MGCP-based, or SIP-based device authentication configuration, and the reliability configuration of the voice service includes the dual-homing configuration and the emergency standalone configuration.

23.29.1 Configuring Device Authentication

Device authentication is a method to improve the security of the core network and prevent illegal devices from registering with the core network device.

Configuring Device Authentication (H.248-based)

This topic describes how to configure the H.248-based device authentication to prevent illegal MGs from registering with the MGC.

Prerequisite

- The MG interface must be configured successfully.
- The parameters, including the encryption type, the initial key and the DH authentication, and the MG ID, must be configured on the MGC. These parameters must be the same as the parameters configured on the MA5600T/MA5603T/MA5608T.

Precautions

If Huawei products such as the SoftX3000 is used as the MGC, the authentication MG ID must be a character string with more than eight bits.

Procedure

In the global config mode, run the **interface h248** command to enter the MG interface mode.

Step 1 Run the **mg-software parameter 4** command to configure the registration mode.

Step 2 Run the **mg-software parameter 6 0** command to configure the device authentication function on the MG interface.

Step 3 Run the **auth** command to configure the authentication MG ID and the initial key.

Step 4 Run the **display auth** command to query the authentication parameters.

Step 5 Run the **reset coldstart** command to reset the MG interface.

Reset the MG interface to make the MG interface register with the MGC (and to make the modified attributes of the MG interface take effect) so that the MG interface can work in the

normal state. The MG interface can be enabled in different ways (see Parameters of the **reset** command). For a newly configured MG interface, enable the MG interface through cold start.

----End

Example

Configure the authentication parameters for the MA5600T/MA5603T/MA5608T as listed in Table 23-59.

Table 23-59 Data plan for configuring the H.248-based authentication

Item	Data
MG ID	0
Whether the wildcard is used in the registration	Yes
Authentication MG ID	MA5600T/MA5603T/MA5608T. It must be the same as the authentication MG ID on the MGC. Otherwise, the MG cannot register with the MGC.
Initial key	0123456789ABCDEF. It must be the same as the initial key configured on the MGC.

The following is a configuration example based on the data plan:

```
huawei(config)#interface h248 0
huawei(config-if-h248-0)#mg-software parameter 4 0
huawei(config-if-h248-0)#display mg-software parameter 4
  -------------------------------------------------
  Interface Id:0         para index:4   value:0
  -------------------------------------------------
 APPENDIX:
  -------------------------------------------------
   Interface software parameter name:
   4: Whether MG register to MGC with wildcard
      0: Yes
      1: No
huawei(config-if-h248-0)#mg-software parameter 6 0
huawei(config-if-h248-0)#display mg-software parameter 6
  -------------------------------------------------
  Interface Id:0         para index:6   value:0
  -------------------------------------------------
 APPENDIX:
  -------------------------------------------------
   Interface software parameter name:
   6: Whether MG support authentication
      0: Yes
      1: No
huawei(config-if-h248-0)#auth auth_mgid MA5600T/MA5603T/MA5608T initial_key
0123456789ABCDEF
huawei(config-if-h248-0)#display auth
```

```
 [AUTH_PARA config]
  Initial Key   : 0123456789ABCDEF
  Auth MGid     : MA5600T/MA5603T/MA5608T
  Algorithm     : MD5
huawei(config-if-h248-0)#reset coldstart
  Are you sure to reset MG interface?(y/n)[n]:y
```

Configuring Device Authentication (MGCP-based)

This topic describes how to configure the MGCP-based authentication parameters for the MG interface on the MA5600T/MA5603T/MA5608T to implement device authentication and prevent illegal MGs from registering with the MGC.

Prerequisite

- The MG interface must be configured successfully.
- The parameters, including the encryption type, the initial key and the DH authentication, and the MG ID, must be configured on the MGC. These parameters must be the same as the parameters configured on the MA5600T/MA5603T/MA5608T.

Procedure

In the global config mode, run the **interface mgcp** command to enter the MG interface mode.

Step 1 Run the **mg-software parameter 4** command to configure the registration mode.

Step 2 Run the **auth** command to configure the authentication MG ID and the initial key.

If Huawei products such as the SoftX3000 is used as the MGC, the authentication MG ID must be a character string with more than eight bits.

NOTE

When the MGCP protocol is used, the MG interface supports two authentication modes:

- Passive authentication mode: In this mode, the device registers with the MGC and is authenticated only after required by the MGC.
- Active authentication mode: In this mode, the device is authenticated when the device registers with the MGC.

In actual applications, you can select the authentication mode according to the requirements.

Step 3 Run the **display auth** command to query the authentication parameters.

Step 4 Run the **reset** command to reset the MG interface.

----End

Example

To configure the authentication parameters for the MA5600T/MA5603T/MA5608T as listed in Table 23-60, do as follows:

Table 23-60 Data plan for configuring the MGCP-based device authentication

Item	Data
MG ID	0

Item	Data
Whether the wildcard is used in the registration	Yes
Authentication mode	Active authentication mode
Authentication MG ID	MA5600T/MA5603T/MA5608T. It must be the same as the authentication MG ID on the MGC. Otherwise, the MG cannot register with the MGC.
Initial key	0123456789ABCDEF. It must be the same as the initial key configured on the MGC.

The following is a configuration example based on the data plan:

```
huawei(config)#interface mgcp 0
huawei(config-if-mgcp-0)#mg-software parameter 4 0
huawei(config-if-mgcp-0)#display mg-software parameter 4
  -------------------------------------------------
  Interface Id:0         para index:4   value:0
  -------------------------------------------------
 APPENDIX:
  -------------------------------------------------
   Interface software parameter name:
   4: Whether MG register to MGC with wildcard
      0: Yes
      1: No
huawei(config-if-mgcp-0)#auth mode2 auth_mgid MA5600T/MA5603T/MA5608T initial_key
0123456789ABCDEF
huawei(config-if-mgcp-0)#display auth
  active request authentication mode config:
  Initial Key    : 0123456789ABCDEF
  Auth MGid      : MA5600T/MA5603T/MA5608T
  Algorithm      : MD5
huawei(config-if-mgcp-0)#reset
  Are you sure to reset MG interface?(y/n)[n]:y
```

Configuring Device Authentication Based on SIP

When the Session Initiation Protocol (SIP) is used, the voice service of the MA5600T/MA5603T/MA5608T supports the authentication for a SIP interface and single user in user name+password or user name+HA1 mode.

Prerequisite

- The SIP interface has been added. For details about how to add a SIP interface, see 23.17.1 Configuring an SIP Interface.
- The authentication information has been configured on the IP multimedia subsystem (IMS) side.

Context

- The device authentication must be supported on the IMS side. Ensure that the authentication data on the device side is the same as that on the IMS side.
- The user authentication on the MA5600T/MA5603T/MA5608T running SIP involves SIP interface authentication and user authentication. In SIP interface authentication, proxy option detection messages are authenticated. In user authentication, user registration and call messages are authenticated.
- A SIP user can be authenticated based on a SIP interface or a single user. Run the **sip-auth parameter auth-mode** command to configure a user authentication mode.
  - If the user authentication mode is set to *interface*, only the user name and password configured based on a SIP interface can be used for user authentication when the user authentication is based on both a SIP interface and a single user.
  - If the user authentication mode is set to *single-user*, the user name and password configured based on a single user are preferentially used for user authentication when the user authentication is based on both a SIP interface and a single user. The default user authentication mode is *single-user*.

Procedure

- Perform the authentication based on a SIP interface.
  a. In the global config mode, run the **interface sip** command to enter the SIP interface mode.
  b. Run the **sip-auth-parameter** command to configure the authentication user name and password for the SIP interface.

     Security authentication information includes password authentication mode, user name, password, and user authentication mode.
     - Password authentication mode includes **password** and **ha1**. In **password** mode, the original user password is configured. In **ha1** mode, a password is generated after the original user password is encrypted by using the message digest 5 (MD5) algorithm.
     - User authentication mode includes **interface** and **single-user**. The **interface** mode indicates that authentication is performed based on interface. This means that all users under an interface share an authentication user name. The **single-user** mode indicates that each user has a unique identity.

  c. Run the **reset** command to reset the SIP interface.
- Perform the authentication based on a single user.
  a. In global config mode, run the **esl user** command to enter extend signaling link (ESL) user mode.
  b. According to the service type, run the **sippstnuser auth set** command or the **sipbrauser auth set** command or the **sipprauser auth set** command to configure the authentication user name, password for single user.
  c. Run the **display sippstnuser authinfo** command or the **display sipbrauser authinfo** command or the **display sipprauser authinfo** command to query the security authentication information.

----End

Example

Configure the security authentication information of SIP interface 0 on the MA5600T/MA5603T/MA5608T, where,

- User authentication mode is **interface**
- Password authentication mode is **password**
- User name is **huawei.com**
- Password is **123456789**

```
huawei(config)#interface sip 0
huawei(config-if-sip-0)#sip-auth-parameter auth-mode interface password-mode pas
sword
  User Name(<=64 characters, "-" indicates deletion):huawei.com
  User Password(<=64 characters, "-" indicates deletion):   //Enter password here.
  The configuration will take effect after resetting the interface
huawei(config-if-sip-0)#reset
  Are you sure to reset the SIP interface?(y/n)[n]:y
```

Configure the security authentication information of the PSTN user on port 0/2/1, where,

- Telephone number is 88810001
- Authentication password mode is **password**
- User name is **huawei**
- Password is **huawei123**

To configure the authentication data of such a PSTN user, do as follows:

```
huawei(config)#esl user
huawei(config-esl-user)#sippstnuser auth set 0/2/1 telno 88810001 password-mode
password
  User Name(<=64 characters, "-" indicates deletion):huawei
  User Password(<=64 characters, "-" indicates deletion):   //Enter password here.
```

23.29.2 Configuring Inner Standalone (H.248-based or SIP-based)

This topic describes how to configure the inner standalone. After the inner standalone is configured, the internal phones can call each other using the internal extension numbers even if the interface between the gateway and the softswitch is interrupted.

Context

- The MG interface supports the inner standalone function only when it uses the H.248 protocol.
- When the MG interface works in the inner standalone state, only the internal users of the MG interface can communicate in the normal state.
- To maintain the same user phone number in the standalone state as the one used in normal condition, configure the phone number on the MG to be the same as that on the MGC.

Prerequisite

- The voice service users are configured properly on the H.248/SIP interface and the users can call each other successfully.
- The user phone number of the H.248 interface/SIP interface is configured to be the same as that on the softswitch. (Command: **mgpstnuser modify** (H.248 interface)/**sippstnuser modify** (SIP interface)).

Procedure

- Configure inner standalone (based on H.248 protocol)
  a. Run the **mg-software parameter 11 1** command to configure the MG interface to support the inner standalone function.
  b. Run the **digitmap set inner** command to configure the internal digitmap.

  NOTE
  The configured digitmap should correspond to the user phone number.

  c. (Optional) Run the **standalone parameters** command to configure the inner standalone timers.
     - The inner standalone timers include the dial tone timer (default: 10s), the busy tone timer (default: 40s), and the ringing tone timer (default: 50s).
     - Generally, the default inner standalone timers can be used.
- Configure inner standalone (based on SIP protocol)
  a. Run the **mg-software parameter 2 1** command to configure the SIP interface to support the inner standalone function.

     By default, the SIP inner standalone function is disabled.
  b. (Optional) Run the **local-digitmap add** command to add a digitmap used in SIP inner standalone.

     The system has a digitmap named **defaultnormal**, which can match any phone numbers. Therefore, a new digitmap does not need to be added unless otherwise required.

----End

Example

Assume that an incoming third-party call will interrupt the inner standalone call after the communication between MG 0 and the MGC recovers. To configure MG 0 to support the inner standalone, and set the internal digitmap to 1234xxxx, do as follows:

```
huawei(config)#interface h248 0
huawei(config-if-h248-0)#mg-software parameter 11 1
huawei(config-if-h248-0)#display mg-software parameter 11
  ------------------------------------------------
  Interface Id:0         para index:11  value:1
  ------------------------------------------------
 APPENDIX:
  ------------------------------------------------
   Interface software parameter name:
   11: Stand alone support
      0: None
      1: Inner
      2: Emergency
```

```
     3: Both
huawei(config-if-h248-0)#digitmap set inner 1234xxxx
huawei(config-if-h248-0)#display digitmap
  ---------------------------------------------------------------------------
  Inner digitmap                                  : 1234xxxx
  Emergency digitmap                               : -
  Urgent digitmap (for overload or bandwidth restrict)   : -
  Dualdial digitmap for card service                 : -
  ---------------------------------------------------------------------------
```

To configure SIP interface 0 to support the inner standalone (so that the internal phones whose numbers belong to the number segment of 07552856xxxx can call each other, if the communication between SIP interface 0 and the softswitch is interrupted), do as follows:

```
huawei(config)#display local-digitmap name defaultnormal

  Command:
        display local-digitmap name defaultnormal
  ----------------------------------------
  Name: defaultnormal
  Type: normal
  Body: x.S|Exx.S      //can match any phone numbers
  ----------------------------------------
huawei(config)#interface sip 0
huawei(config-if-sip-0)#mg-software parameter 2 1
huawei(config-if-sip-0)#display mg-software parameter 2
  ------------------------------------------------
  MGID:0         para index:2   value:1
  ------------------------------------------------
 APPENDIX:
  ------------------------------------------------
 Parameter Index:  Interface software parameter name:
   2 : SAL Support
       0: No
       1: Yes
```

23.29.3 Configuring the Dual Homing (Multi-Homing)

This topic describes how to configure the H.248-based, MGCP-based, and SIP-based dual homing (multi-homing). Dual homing (multi-homing) is a measure that protects the softswitch against a crash and a disaster recovery mechanism against accidents (such as a fire in the telecommunications room, disconnection of the cable connected to the telecommunications room, and abnormal power supply).

Context

The working principle of multi-homing is as follows: an MG interface (for H.248 and MGCP) or a SIP interface is homed to multiple registration servers. If the primary server malfunctions, the interface is switched to a secondary server and then continues to provide services.

Dual homing is one type of multi-homing configured with only primary/secondary servers without a disaster-recovery server.

Configuring H.248-based Dual Homing (Multi-homing)

In the case of H.248, the MA5600T/MA5603T/MA5608T supports homing of a media gateway (MG) interface to the primary/secondary media gateway controllers (MGCs) and disaster-recovery MGC. When the primary MGC malfunctions, the MG interface will register with the secondary MGC and then the disaster-recovery MGC cyclically.

Context

Technically speaking, dual homing is a configuration in which an MG is homed to the primary MGC and secondary MGC. Multi-homing is a configuration in which an MG is homed to the primary MGC, secondary MGC, and disaster-recovery MGC. Multi-homing is an enhancement of dual homing. In a broad sense, dual homing is one type of multi-homing.

The MA5600T/MA5603T/MA5608T provides different application policies for the dual homing (multi-homing) by configuring MG homing parameters and MG software parameters.

Procedure

Create an MG interface and configure MG interface parameters.

1. In the global config mode, run the **interface h248** command to enter the MG interface mode.
2. Run the **if-h248 attribute** command to create an MG interface and then configure the primary/secondary MGCs and disaster-recovery MGC.

   When configuring an MG interface supporting dual homing, note that:

   - The MG is dual homed to the primary/secondary MGCs. The configurable parameters include **secondary-mgc-ip1** *secondary-mgc-ip1*, **secondary-mgc-ip2** *secondary-mgc-ip2*, **secondary-mgc-port** *secondary-mgc-port*, or **mgc-domain-name2** *mgcdomainname2*.
   - At least one secondary MGC (containing the IP address and port ID) is configured.

   When configuring an MG interface supporting multi-homing, note that:

   - A disaster-recovery MGC is configured based on the dual homing. The configurable parameters include **stand-alone-mgc-ip1** *stand-alone-mgc-ip1*, **stand-alone-mgc-ip2** *stand-alone-mgc-ip2*, and **stand-alone-mgc-port** *stand-alone-mgc-port*.
   - At least one disaster-recovery MGC (containing the IP address and port ID) is configured.

   For details about how to configure an MG interface, see 23.20.1 Configuring an MG Interface.

Step 1 Configure the software parameters of an MG interface.

1. Run the **mg-software parameter 2** command to configure the MG interface supporting dual homing.

   The values of **mg-software parameter 2** are described as follows:

   - When the value of **mg-software parameter 2** is **0** (default value), multi-homing is not supported. Specifically, after an MG interface is unable to register with the primary MGC, the MG interface does no initiate registration with the secondary MGC or disaster-recovery MGC though it has been configured.
   - When the value of **mg-software parameter 2** is **1**, multi-homing is supported but auto-switching is not supported. Specifically,

- An MG interface registering with the primary MGC will register with the secondary MGC and then the disaster-recovery MGC cyclically when the primary MGC malfunctions.
- After the primary MGC recovers, the secondary MGC or disaster-recovery MGC will not automatically switch back to the primary MGC. Run the **mgc switch(h248)** command to forcibly switch the MGC to the primary MGC.

– When the value of **mg-software parameter 2** is **2**, multi-homing and auto-switching are supported. Specifically,

- An MG interface registering with the secondary MGC will automatically switch to the primary MGC after the primary MGC recovers.
- An MG interface registering with the disaster-recovery MGC will automatically switch to the primary or secondary MGC after the primary or secondary MGC recovers.

NOTICE

After an MG interface is manually or automatically switched, the MG interface will restart, causing services interrupted for a short time.

2. (Optional) Run the **mg-software parameter 36** command to configure the registration interval during MGC's multi-homing registration switchover.

   This parameter is used to configure the switch interval when an MGC (primary MGC, secondary MGC, or disaster-recovery MGC) malfunctions and switches to another MGC. It is defaulted to 40s.

3. Run the **display mg-software parameter** command to query the software parameters of the MG interface.

----End

Example

To configure dual homing on MG interface 0 with the following settings, do as follows:

- The IP address of the primary MGC is 192.168.0.10 and the port number for the transport layer protocol is 2944.
- The IP address of the secondary MGC is 192.168.0.20 and the port number for the transport layer protocol is 2944.
- Auto-switching is not supported.

```
huawei(config)#interface h248 0
huawei(config-if-h248-0)#if-h248 attribute primary-mgc-ip1 192.168.0.10 primary-mgc-port 2944 secondary-mgc-ip1 192.168.0.20 secondary-mgc-port 2944
huawei(config-if-h248-0)#mg-software parameter 2 1
huawei(config-if-h248-0)#display mg-software parameter 2
  ------------------------------------------------
  Interface Id:0         para index:2   value:1
  ------------------------------------------------
 APPENDIX:
  ------------------------------------------------
   Interface software parameter name:
   2: Whether MG support multi-home function
```

```
    0: Do not support the multi-homing
    1: Support the multi-homing, but do not support the auto switchover
    2: Support the multi-homing and auto switchover
```

To configure dual homing on MG interface 1 with the following settings, do as follows:

- The IP address of the primary MGC is 192.168.0.10 and the port number for the transport layer protocol is 2944;
- The IP address of the secondary MGC is 192.168.0.20 and the port number for the transport layer protocol is 2944;
- The IP address of the disaster-recovery MGC is 192.168.0.30 and the port number for the transport layer protocol is 2945;
- Auto-switching is supported.

```
huawei(config)#interface h248 1
huawei(config-if-h248-1)#if-h248 attribute primary-mgc-ip1 192.168.0.10 primary-
mgc-port 2944 secondary-mgc-ip1 192.168.0.20 secondary-mgc-port 2944 stand-alone
-mgc-ip1 192.168.1.30 stand-alone-mgc-port 2945
huawei(config-if-h248-1)#mg-software parameter 2 2
huawei(config-if-h248-1)#display mg-software parameter 2
  -------------------------------------------------
  Interface Id:1         para index:2   value:2
  -------------------------------------------------
 APPENDIX:
  -------------------------------------------------
   Interface software parameter name:
   2: Whether MG support multi-home function
      0: Do not support the multi-homing
      1: Support the multi-homing, but do not support the auto switchover
      2: Support the multi-homing and auto switchover
```

Configuring MGCP-based Dual Homing

This topic describes how to configure the MGCP-based dual homing.

Context

The MA5600T/MA5603T/MA5608T supports registering with three MGCs (MGC1, MGC2, and MGC3) through the MG interface. MGC1 serves as the primary MGC. When MGC1 fails, the MG can switch to MGC2 and continue working. When MGC2 also fails, the MG can switch to MGC3 and continue working.

Prerequisite

- MGC1 and MGC2 must be configured in the attributes of the MG interface.
- On the MGCs, the data for interconnecting with the MG must be configured.

Procedure

In the global config mode, run the **interface mgcp** command to enter the MG interface mode.

Step 1 Run the **mg-software parameter 3 1** command to enable the heartbeat message function.

Step 2 Run the **mg-software parameter 2 0** command to configure the MG interface to support dual homing.

----End

Example

To configure MG interface 0 to support dual homing and enable the heartbeat message function, do as follows:

```
huawei(config)#interface mgcp 0
huawei(config-if-mgcp-0)#mg-software parameter 3 1
huawei(config-if-mgcp-0)#mg-software parameter 2 0
```

Configuring the SIP-based Dual Homing

the MA5600T/MA5603T/MA5608T supports the 1+1 mutual assistance mode (the active/standby mode) of the upstream proxy devices. When either of the upstream active/standby devices is faulty, the MA5600T/MA5603T/MA5608T automatically switches the service to the other device. In this way, the disaster recovery solution is implemented through SIP to improve the access reliability of the device.

Context

The MA5600T/MA5603T/MA5608T supports the SIP interface homing to two proxy servers (Proxy1 and Proxy2), where Proxy1 functions as the primary proxy server. When Proxy1 fails, the MG can switch to Proxy2 to continue working.

Prerequisite

- The data for interconnecting with the SIP interface must be configured on the IMS.
- When configuring the IP address of the SIP interface, make sure that the IP address (signaling IP address or media IP address) exists in the corresponding IP address pool.

Procedure

In the global config mode, run the **interface sip** command to enter the SIP interface mode.

Step 1 Run the **if-sip attribute basic** command to configure the basic attributes of the SIP interface.

To support dual homing, the information about the secondary proxy server must be specified here. A proxy server can be identified by its IP address or domain name.

Step 2 Run the **reset** command to reset the interface.

----End

Example

Assume that the SIP interface ID is 0, the media IP address is 10.10.10.13, signaling IP address is 10.10.10.13, transfer protocol is UDP, and UDP port number is 5000; IP address 1 of the primary proxy server is 10.10.10.14, and port number of the primary proxy server is 5060; IP address 1 of the secondary proxy server is 10.10.10.15, and port number of the secondary proxy server is 5060; the homing domain name **huawei.com**, and profile ID is 1. To configure the dual homing attributes of SIP interface 0, do as follows:

```
huawei(config)#interface sip 0
huawei(config-if-sip-0)#if-sip attribute basic media-ip 10.10.10.13
signal-ip 10.10.10.13 signal-port 5000 transfer udp primary-proxy-ip1
10.10.10.14 primary-proxy-port 5060 primary-proxy-domain proxy.domain
secondary-proxy-ip1 10.10.10.15 secondary-proxy-port 5060
huawei(config-if-sip-0)#if-sip attribute basic home-domain huawei.com sipprofile
-index 1
huawei(config-if-sip-0)#reset
```

24 Device Management

About This Chapter

This topic covers the overview, general specifications, availability, and sub-features of device management security.

24.1 Introduction

Device security includes the following features: SNMP, inband management VPN, SSH, user management, remote connection security, log management, version and data management, and alarm and event management.

Feature	Description
SNMP	The NMS communicates with the device through SNMP.
Inband management VPN	The carriers use the virtual private network (VPN) to manage and maintain devices and the management protocol on the device can use virtual routers for route forwarding.
SSH	Based on the application layer and transport layer, SSH is a protocol that provides security for remote login session and other network services. It is used for remote management connection and file transfer.
User management	User management involves management of user rights and encryption of the user name and password.
Remote connection security	It includes a series of firewall functions aiming at users' login connection to the device, and the function of disabling the device service port.
Log management	Logs include the security event logs relevant to the system security events and the operation logs of users.
Version and data management	This management function includes patch management, rollback function, configuration data management, and version upgrade.
Alarm and event management	This management function includes recording and setting alarms and events and collecting their statistics.

24.2 Remote Operation

Remote operation refers to performing routine maintenance on the device remotely, without any on-site visit. This feature greatly reduces operating expenditures (OPEX) of carriers and improves user satisfaction. If a user service is abnormal, the carrier can remotely identify the fault cause and restore the service within shot time.

Remote operation supports outband telnet and inband telnet.

- Outband telnet

  The interface used by outband telnet is the only one Ethernet port (RJ-45) on the front panel of the control board. After the IP address and related route are configured on this port, the device can be logged in to through this port in the telnet mode for related operations and maintenance.
- Inband telnet
  - The interface used by inband telnet is the VLAN Layer 3 interface inside the device. The system supports up to 32 IP addresses for the VLAN Layer 3 interface and the subnets of these IP addresses must be different.
  - In the case of remote telnet, it is recommended to configure the acceptable/refused IP address segments to prevent the login of a user who uses an illegal IP address.

24.3 SNMP

This topic provides an introduction to the SNMP sub feature, and then describes the working principle of this sub feature.

24.3.1 Introduction

Definition

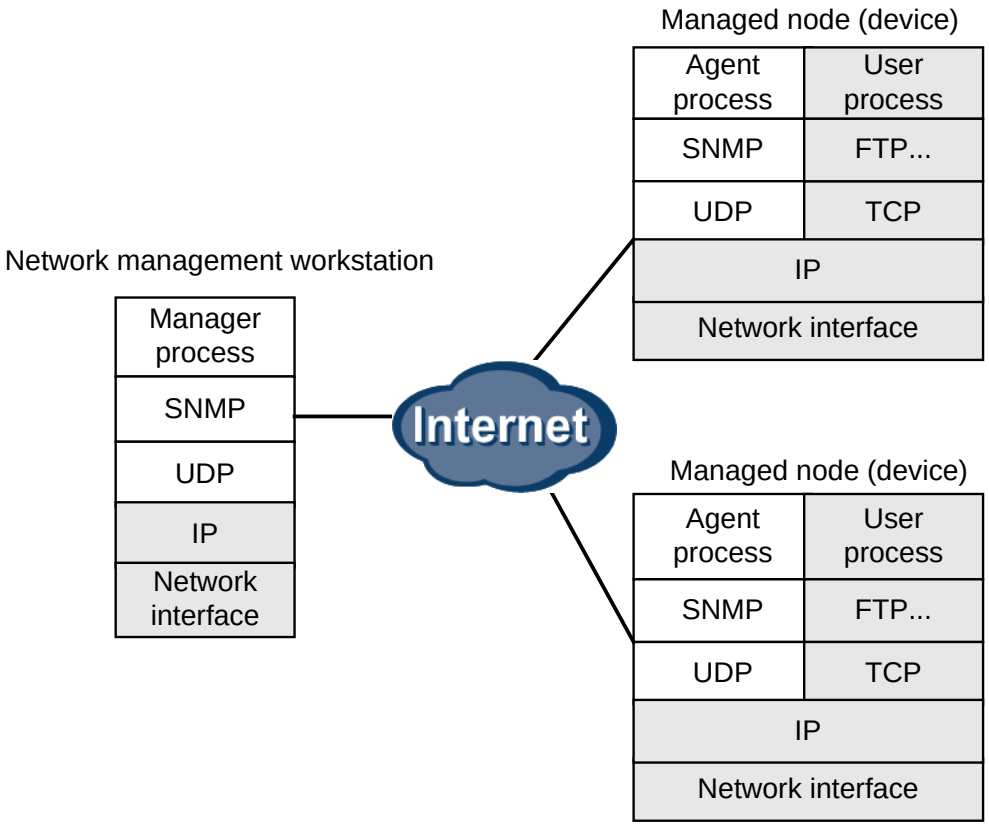
The Simple Network Management Protocol (SNMP) is a network management protocol that is widely used in the TCP/IP network. It provides a means of managing network resources using a central computer (network management workstation) that runs the network management software.

Network management involves four parts:

- Managed node: device that is monitored, namely NE.
- Agent: software used to display the status of the managed nodes (devices).
- Network management workstation: central device that communicates with the agents of the managed nodes and displays the status of the agents.
- Network management protocol: protocol (such as SNMP) for information exchange between the network management workstation and the agent.

Figure 24-1 shows the typical configuration of an SNMP-managed network. The entire network must have at least one network management workstation, which acts as the network management center and runs the manager process. Each managed node must have an agent. The manager and the agent communicate with each other using UDP-based SNMP messages.

Figure 24-1 Typical configuration of an SNMP-managed network



Purpose

SNMP is mainly used for network management. There are two types of network management, as described in the following:

- One is management of network applications, user account, and access right (permission). Such management is related to software and is not described in detail.
- The other is management of NEs such as the MA5600T/MA5603T/MA5608T. Generally, the managed devices are far away from the central telecommunications room where the network management engineers work. When such devices are faulty, it is ideal if the network management engineers are notified of the faults automatically. However, devices such as the MA5600T/MA5603T/MA5608T cannot do the same as users making phone calls to notify the network management engineers of its application faults.

To resolve such an issue, equipment vendors provide network management functions for some devices. In this way, the network management workstation can query the status of managed devices remotely; likewise, the managed devices send alarms to the network management workstation when events of a specific type occur.

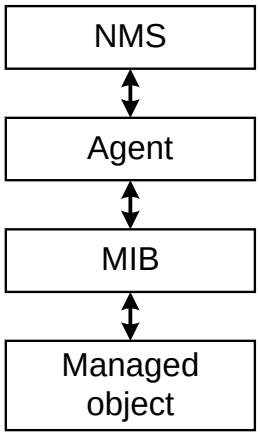
24.3.2 SNMP Network Management Model

Through the SNMP protocol, signaling is exchanged between the network management workstation and the agent.

- The manager in the network management workstation sends an SNMP request PDU to the agent.
- After obtaining the required information following the query of the MIB of managed devices, the agent sends an SNMP response PDU to the manager.
- When the managed device is abnormal, the agent notifies the manager of the fault through a trap, which helps the network management engineers solve the problems in time.

Figure 24-2 shows the SNMP network management model.

Figure 24-2 SNMP network management model



Implementation of SNMP network management consists of three parts: management information base (MIB), structure of system management (SMI), and SNMP.

24.3.3 SNMP MIB

The management information base (MIB) is an abstract set of all managed objects. MIB is tree-structured and therefore is called the MIB tree. Each managed object corresponds to a leaf in the MIB tree and is called a MIB leaf.

The MIB tree is a static tree, that is, the MIB tree structure completes initialization after the device is started. After that, the manager only searches for or modifies the contents of each managed object. The manager manages devices by reading information from and writing information to the managed objects in the MIB.

24.3.4 SNMP SMI

The structure of management information (SMI) defines a set of rules of naming and defining managed objects to achieve communication between SNMP entities.

SNMP is a protocol running at the application layer, which requires the protocol entities at the two ends to exchange PDUs. However, data at the lower layer is byte sequence. In this case, SMI is applied to help SNMP protocol entities to change the received byte sequence to a PDU and then change the PDU with the internal data structure to a byte sequence that can be sent.

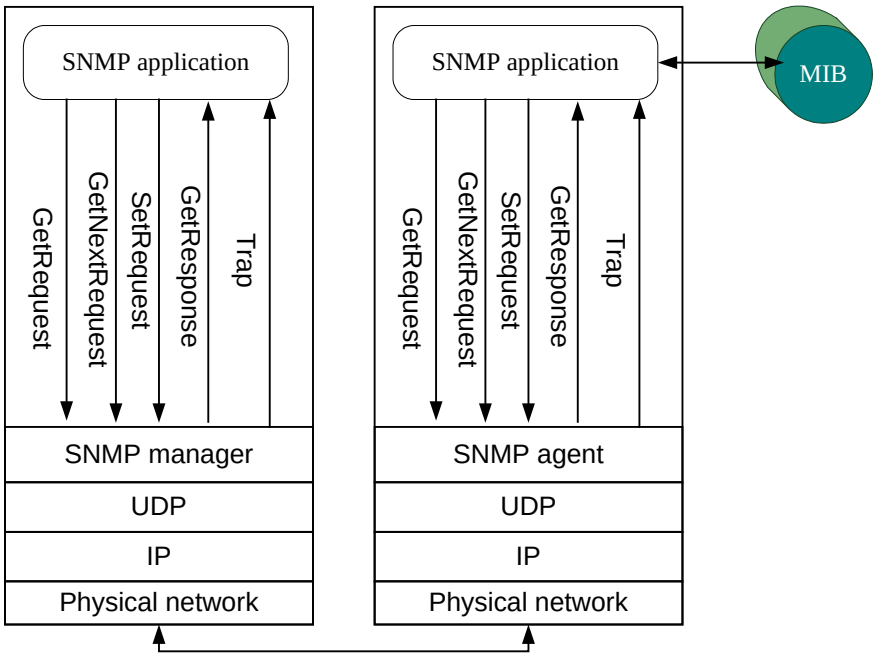
24.3.5 Working Principle of SNMPv1

SNMPv1 specifies five core protocol data units (PDUs), that is, SNMP messages, which are exchanged between the manager and the agent.

- Get-request: Retrieves the value of one or more parameters from the agent.
- Get-next-request: Retrieves the value of the next parameter from the agent lexicographically.
- Set-request: Sets the value of one or more parameters for the agent.
- Get-response: Returns the value of one or more parameters. This operation is sent by the agent and is a response to the preceding three operations.
- Trap: PDU sent actively by the agent to notify the manager of the occurrence of certain events. When a device generates an alarm indicating that important data of the device is changed by the user, console, or another manager, the agent notifies the manager of such information through traps. After receiving the traps, the manager generates relevant actions (such as polling) to diagnose faults.

The first three operations are sent from the manager to the agent and the last two from the agent to the manager, as shown in Figure 24-3.

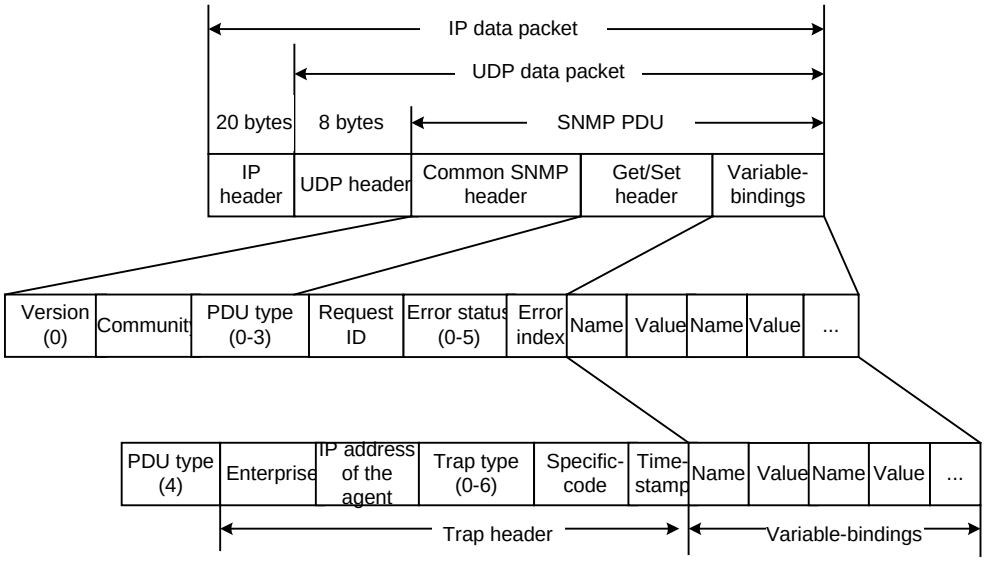
Figure 24-3 Exchange of SNMPv1 PDUs



SNMP PDU Structure

An SNMP PDU consists of the common SNMP header, get/set header, trap header, and variable bindings, as shown in Figure 24-4.

Figure 24-4 SNMP PDU structure



- Common SNMP header

  The common SNMP header consists of three fields:

  - Version. The value of this field is the PDU version minus one. For example, the value of this field for the SNMPv1 PDU is 0.
  - Community. It is the password in plain text used between the manager and the agent, in the format of character string. A common community name is public, a string of six characters.
  - PDU type. There are five types of PDU, as listed in Table 24-1.

Table 24-1 SNMP PDU type

PDU Type	Name
0	Get-request
1	Get-next-request
2	Get-response
3	Set-request
4	Trap

- Get/Set header
  - Request ID

    It is an integer set by the manager. When sending the get-response PDUs, the agent also needs to return the request ID. The manager can send the get PDUs to multiple agents using the UDP port. However, the response PDU for the first get PDU does not necessarily arrive first. Considering such a situation, the request ID is set so that the manager can correlate incoming response PDUs with corresponding request PDUs.

- Error status

  It is filled when the agent responds to the manager, as described in Table 24-2.

Table 24-2 Error status

PDU Type	Name	Description
0	noError	No error occurs.
1	tooBig	The agent fails to put the response into an SNMP PDU.
2	noSuchName	The operation specifies a non-existent variable.
3	badValue	A set operation specifies an invalid value or syntax.
4	readOnly	The manager is trying to modify a read-only variable.
5	genErr	Some other errors occur.

- Error index

  When an error such as noSuchName, badValue, or readOnly occurs, the agent sets an integer as the error index during its response. The error index specifies the position of the error variable in the variable list.

- Trap header
  - Enterprise

    This field is filled with the object ID of the network device carried in the trap PDU.
  - Trap type

    The formal name of this field is generic-trap. There are seven trap types, as described in Table 24-3.

Table 24-3 Trap type

Trap Type	Name	Description
0	coldStart	The agent is initialized.
1	warmStart	The agent is re-initialized.
2	linkDown	A port changes from the working state to the faulty state.
3	linkUp	A port changes from the faulty state to the working state.
4	authenticationFaliure	A PDU with an invalid community name is received from the SNMP manager.
5	egpNeighborLoss	The MA5600T/MA5603T/MA5608T, an EGP neighbor, is faulty.
6	enterpriseSpecific	This field indicates an event defined by the

Trap Type	Name	Description
		agent, which is specified by the specific-code field.

In the case of 2, 3, or 5, the first variable in the variable-bindings of a PDU needs to specify the port that is used for response.

- Specific-code

  This field specifies the event (for example, trap type 6) defined by the agent. If the event is not defined by the agent, this field is filled with 0.

- Time stamp

  This field specifies the time elapsed between the initialization of the agent and the generation of the trap, in the unit of 10 ms. For example, if the time stamp is 1908, it indicates that the trap is generated 19080 ms after the initialization of the agent.

- Variable-bindings

  This field specifies the name and value of one or more variables. In the get or get-next PDU, this field is filled with 0.

24.3.6 Working Principle of SNMPv2c

Simplicity is a key to SNMP success, which caters to the need of clear management protocols in a large-size and complicated network involving devices of multiple vendors. However, to achieve simplicity, SNMP sacrifices certain functions, for example:

- SNMP does not provide the bulk access mechanism, causing low access efficiency of large data.
- SNMP runs over only TCP/IP. It does not support other network protocols.
- SNMP does not provide the mechanism for communication between managers. It is applicable to centralized management, but not distributed management.
- SNMP can be used for monitoring network devices, but not for monitoring the network.

Aiming at resolving these problems, IETF continuously optimizes SNMP and finally formulates SNMPv2c. SNMPv2c has the following enhancements to SNMPv1:

- Supports new types of PDUs.
- Extends the types supported by SMI.
- Supports communication between managers.

New PDU in SNMPv2c

- GetBulk

  GetBulk is an extension of get-next. That is, a getBulk operation equals multiple get-next operations. With one getBulk operation, a large amount of information can be obtained, which effectively reduces communications between the manager and the agent and thus improves network performance.

24.3.7 Working Principle of SNMPv3

The structure of SNMPv3 is model-based, which facilitates addition and modification of the protocol functions. SNMPv3 has the following advantages:

- Good adaptability: SNMPv3 is applicable to multiple operation environments. It can manage both simple networks and complicated networks.
- Excellent scalability: New models can be added according to actual requirements.
- High security: SNMPv3 provides multiple security processing models.

SNMPv3 has four major models: message processing and control model, local processing model, user-based security model (USM), and view-based access control model (VACM).

Different from SNMPv1 and SNMPv2, SNMPv3 implements access control, identity authentication, and encryption through its local processing model and USM.

Message Processing and Control Model

Defined in RFC2272, the message processing and control model is responsible for generating and analyzing SNMP PDUs and determining whether PDUs need to pass the agent server during transmission. During the generation of a PDU, this model receives the PDU from the dispatcher, and then the USM adds the security parameters to the PDU header. When analyzing a received PDU, the USM processes the security parameters in the PDU header and sends the processed PDU to the dispatcher for processing.

Local Processing Model

The local processing model is mainly used for access control. Access control is to set the information about an agent so that different managers in the management workstation have different rights when accessing the agent. It is implemented through the PDU. Access control can be implemented using the following two methods: by limiting the commands that the manager sends to the agent or by determining the information in the MIB of the agent that the manager visits. The access control method must be set beforehand. SNMPv3 can flexibly determine the access control method through the primitives carrying different parameters.

USM

The USM provides identity authentication and data encryption services. To implement such functions, the manager and the agent must share the same key.

- Identity authentication: When receiving a message, the agent (manager) must determine whether the message is sent from the authorized manager (agent) and whether the message is changed during transmission. This is called identity authentication. RFC2104 defines HMAC, which is an effective tool of generating message authentication codes using cryptographic hash functions and keys. It is widely applied in the Internet. HMAC used by SNMP are HMAC-MD5-96 and HMACSHA-96. HMAC-MD5-96 adopts the MD5 hash function, with the 128-bit authKey as its input. HMACSHA-96 adopts the SHA hash function, with the 160-bit authKey as its input.
- Encryption: It adopts CBC-DES, with the 128-bit privKey as its input. The manager uses a key to calculate the authentication code and then adds the authentication code to the message. After receiving the message, the agent uses the same key to obtain the authentication code and thus decrypts the message. Similar to identity authentication, encryption also requires that the manager and the agent share the same key for message encryption and decryption.

VACM

The VACM implements view-based access control over user groups or community names. A user must first configure a view with rights specified. Then, the user loads the view when

configuring a user, user group, or community name so that the read operation, write operation, or traps (v3) can be limited.

24.3.8 Comparison Between SNMP Protocols in Security

Table 24-4 describes the comparison between SNMP protocols in security.

Table 24-4 Comparison between SNMP protocols in security

SNMP Version	User Authentication	Encryption	Authorization
v1	No; use the community name.	No	No
v2c	No; use the community name.	No	No
v3	Yes; encryption/decryption based on the user name.	Yes	Yes

SNMPv3 USM

SNMPv1 and SNMPv2c lack a security mechanism. SNMPv3 supports the user-based security model (USM) against illegal modification of information and masquerade.

USM mainly checks whether the SNMP message is modified during the network transmission and whether the SNMP message is sent by the alleged user, monitors the outdated SNMP message, and provides the privacy mechanism for SNMP messages.

USM consists of three modules:

- Authentication module: Authenticates the data origin.
- Timeliness module: Prevents message delay or replay.
- Privacy module: Prevents message disclosure.

SNMPv3 VACM

The access control subsystem of the SNMP engine checks whether an access to a special object is allowed. View-based access control model (VACM) is a default access control model in SNMPv3. Compared with SNMPv1 and SNMPv2c, SNMPv3 adopts a more rigorous and dynamic access control model, which facilitates configuration by network management engineers. VACM consists of the following parts:

- Groups

  A group is a set of zero or multiple mappings. It defines all the access rights to all securityNames that belongs to the group. Security level. Different access rights are defined by different security levels.
- Contexts

  An SNMP context is a collection of management information accessible by an SNMP entity.

- MIB views and view families
- Access policy
  - Read-view
  - Write-view
  - Notify-view

24.4 Inband Management VPN

Inband management VPN is a means by which carriers use the virtual private network (VPN) to manage and maintain devices and the management protocol on the device can use virtual routers for route forwarding.

24.4.1 Introduction

Definition

In inband management VPN, associated inband management protocols on the device support the specified VPN instances so that management packets can be received and forwarded using multiple virtual routes. In this way, carriers can use the private network IP address to remotely manage and maintain devices. This method saves public network IP addresses and isolates the management network from the public network.

To achieve inband management VPN, both the inband management server and client must be able to receive the connection requests and data packets from VPN.

Constraints

- The outband management interface belongs to the public network but not VPN. Therefore, only the inband interface but not the outband management interface supports VPN management.
- The servers that can receive VPN requests include:
  - Telnet server
  - SSH server
  - SNMP AGENT (currently, only IPv4 but not IPv6 is supported)
  - TRACE server

NOTE
SSH server is recommended.

- The clients that can receive VPN requests include:
  - FTP client
  - TFTP client
  - SFTP client
  - SNMP TRAP
  - SYSLOG
  - Telnet client

NOTE
SFTP client is recommended.

24.4.2 Principles

Basic Concepts

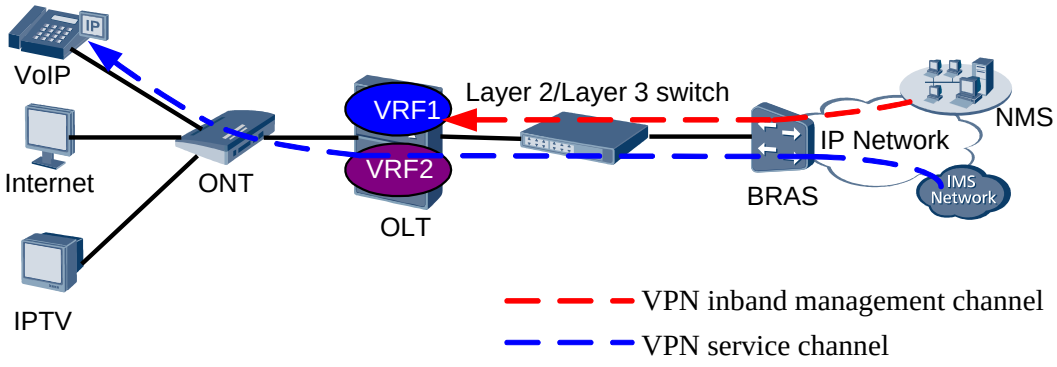
Virtual private network (VPN) is a network technology for encapsulating or encrypting private data and then transmitting the data over the public network. With this technology, the security level of the private network can be provided for the transmitted data and a private network can be constructed based on the public network. VPN is a logical private network that provides the functions of the private network. The network itself, however, is not an independent physical network. In the IP bearer network, VPN is an important measure for logically isolating services, preventing attacks, and helping implement QoS control.

A VPN instance is also called a VPN routing and forwarding table (VRF). Each router is logically divided into multiple virtual routers, that is, multiple VRFs. Each VRF corresponds to a VPN and has its own routing table, forwarding table, and interface. In other words, a router shared by various VPNs is simulated as multiple private routers, thereby isolating VPN routes. Devices that are grouped into a private route exchange routing information of only the private route. After VPN configured successfully, the management packets in the system can be received or sent over the public network.

Inband Management VPN

Inband management VPN uses the VRF function to add the remote NMS and the OLT to the same VPN while on the OLT classifying the management address and VoIP address to different VRFs. In this way, carriers can use the private network IP address to remotely manage and maintain devices. This method saves public network IP addresses and isolates the management network from the public network.

Figure 24-5 Network diagram of inband management VPN



In Figure 24-5, two VRFs (VRF1 and VRF2) are defined on the OLT. VRF2 is the VoIP service channel and VRF1 is the VPN inband management channel.

On the OLT, if a system-level VPN instance for management and maintenance is configured, all related management protocols by default use this VPN instance to connect to or send data to the remote server. In addition, if a user specifies the VPN instance of a single server when configuring the trap destination server or running the telnet command, these two management protocols (for trap and telnet) do not use the system-level VPN instance but instead use the specified VPN instance to connect to the remote server.

24.5 SSH

This topic provides an introduction to the SSH, and describes the working principle of this sub feature.

24.5.1 Introduction

Definition

Secure Shell (SSH) is formulated by the IETF Network Working Group. Based on the application layer and transport layer, SSH provides security for remote login session and other network services.

Purpose

Conventional network service programs such as FTP and telnet transmit password and data in plain text over the network. Unlike these conventional programs, SSH encrypts data to be transferred, which effectively avoids information divulge during remote management. So, it is recommended to use the SSH. In addition, during SSH encryption, data is compressed to a smaller size, which helps achieve faster data transfer.

24.5.2 SSH Working Principle

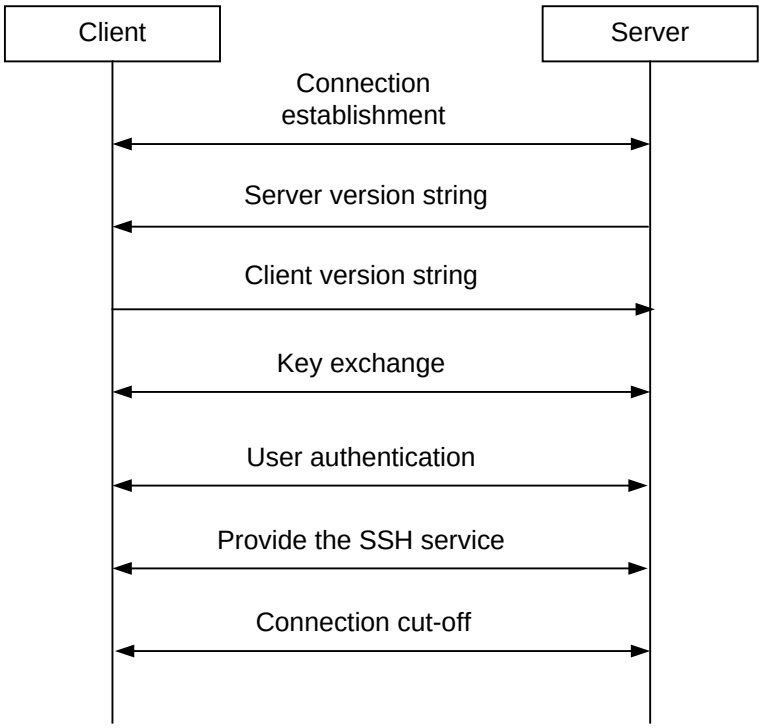
The SSH protocol involves the server and the client.

- As a service daemon, the server responds to connection requests from the client and processes remote connections, including shared key authentication, key exchange, asymmetric encryption, and non-secure connection.
- The client contains the SSH program and applications such as slogin and SFTP. In terms of the client, SSH provides the following two levels of security authentication.
  - One is password-based security authentication. Users can log in to a remote device as long as they know the user name and password for login. In this authentication mode, all data to be transferred is encrypted, but the server to which users are connecting is not always the desired server. That is, maybe some other server pretends to be the desired one.
  - The other is key-based security authentication. In this authentication mode, a pair of keys (server key and host key) need to be created, and the server key needs to be put into the desired server. If a client needs to connect to an SSH server, the client sends a request to the server for security authentication using the host key. Upon receiving the request, the server compares its saved server key with the host key sent by the client. If the two keys are identical, the server sends a "challenge" message encrypted with the server key to the client. After receiving the "challenge" message, the client decrypts the message using the host key and then sends the message back to the server. Till now, the client passes the authentication.

As a security protocol, SSH provides only secure channels but does not transfer data. Through the steps including version negotiation, key exchange, algorithm negotiation, and user authentication, an SSH secure channel is set up. Any data transfer protocol can transfer data in the channel. The tool used by the secure maintenance terminal provides the SSH client function.

Figure 24-6 shows the interaction process between the client and the server using SSH.

Figure 24-6 Interaction process using SSH



24.5.3 SSH-based Encryption for Remote Management Connection

The system supports management of remote operations in the outband or inband telnet mode.

- The port used by outband telnet is the only Ethernet port (RJ-45) on the front panel of the control board. After configuring the IP address and related routes of this port, users can log in to the device through telnet for related operation, maintenance, and management.
- The port used by inband telnet is the VLAN interface of the device. The system supports a maximum of 32 IP addresses for the VLAN interfaces and the subnets of these IP addresses must be different.

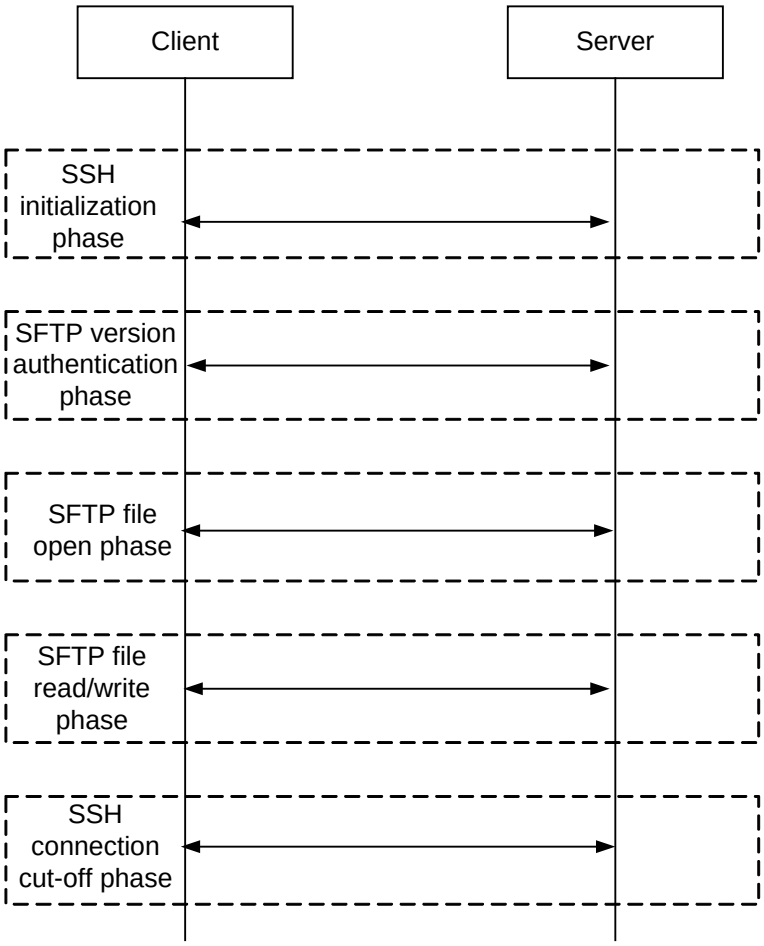
In implementing remote operations, both the secure maintenance terminal and the common maintenance terminal transfer data through telnet. The difference is that the secure maintenance terminal, before transferring data through telnet, encrypts data using SSH. With SSH-based encryption, all the operations are secure after the user logs in to the device through a remote terminal for maintenance and management.

24.5.4 SSH-based Encryption for File Transfer

SFTP is an SSH-based secure file transfer protocol. When a user is authenticated in the password mode, both the user name and password are required on the client. If the user name and password cannot be obtained, file transfer fails.

Figure 24-7 shows the process of file transfer through SFTP.

Figure 24-7 Process of file transfer through SFTP



The process of uploading a file through SFTP is as follows:

1. The client opens the local file that needs to be uploaded to the server.
2. The client sends a request to the server for opening the file on the server.
3. The client writes the local data to the server according to the returned file handle.

Downloading a file through SFTP is based on the SSH authentication:

1. The server and the client both verify the SFTP version in the SFTP stage.
2. The client opens the local file and the remote file.
3. The client reads the corresponding data.
4. The client closes the opened files after reading the data.

24.6 User Management

This topic covers the overview and working principle of user management.

24.6.1 Introduction

Definition

User management involves the following two parts:

- A user needs to be authenticated with user name and password when the user attempts to log in to the device through the command line interface (CLI).
- Users are classified into four levels, namely, super user, administrator, operator, and user. Different levels of users are assigned different operation rights.

Purpose

User management is to ensure the security of device management and maintenance by user name+password authentication and hierarchical right-based management.

24.6.2 Principle

When a user logs in to the system through the CLI, the user must enter the user name and password for authentication. In this way, the user is authenticated to ensure the system security.

Users are classified into four levels, namely, super user, administrator, operator, and user. Different levels of users are assigned different operation rights.

The super user and the administrator have the right to add a user at a lower level, that is:

- The super user can add an administrator, operator, or user.
- The administrator can add only an operator or user.

The system also supports management of user profiles. A user profile supports setting of the following parameters:

- Minimum length of a user name (6-15 characters)
- Minimum length of a password (6-15 characters)
- Validity period of a user name (0-999 days)
- Validity period of a password (0-999 days)
- Start time of user login in the format of hh:mm (for example, 08:30)
- End time of user login in the format of hh:mm (for example, 18:30)

If the validity period of the user name or password is set to 0, it indicates that there is no restriction on the validity period of the user name or password. It is also true for the start time and end time of user login. If other values are set, the user login time is restricted based on the preset values. The system reminds the user through a message three days before the user name and password expire.

After the preceding settings, the security of system management is enhanced to a certain extent. When created, if a user is bound to a user profile and the start time of user login in the user profile is set to 08:30, it indicates that the user cannot log in to the system before 08:30. After a user profile is set, the user profile can be directly bound to a user when adding the user. In addition, the user profile bound to the user that is already created can be modified. A user supports a maximum of 12 user profiles.

The system provides four default user profiles named **root**, **admin**, operator, and **commonuser**, which helps manage and create users in a unified way.

Different names of user profiles indicate the differences in the preceding security settings for the user profiles rather than the differences in user levels. The user level is specified when a user is added. In a root profile, restrictions on users are disabled so that the user bound to the profile can log in to the system after upgrade. It is not recommended that this profile be bound when adding a user.

24.7 ANCP

The Access Node Control Protocol (ANCP) is used by the broadband network gateway (BNG) to manage the line parameters (including QoS and user) of the access node (AN).

24.7.1 Introduction

Definition

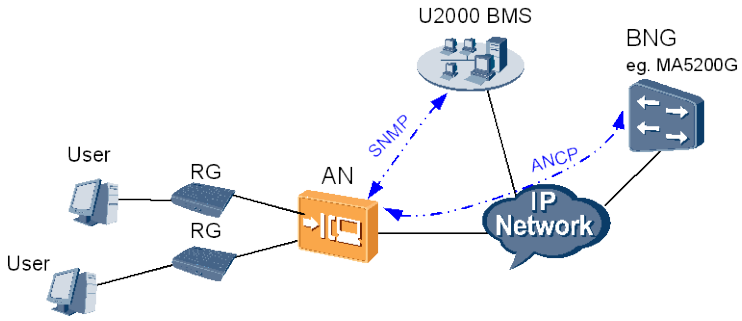
The Access Node Control Protocol (ANCP) is used by the broadband network gateway (BNG) to manage the line parameters (including QoS and user) of the access node (AN).

NOTE
A BNG can be a BRAS or a router, such as the MA5200G.

Figure 24-8 displays the NEs relevant to ANCP.

- The user powers on, disables, or connects the RG to change the line status.
- The BNG and the AN exchange ANCP messages.
- The network administrator manages the AN through the N2000 BMS by using SNMP.

Figure 24-8 ANCP network topology



Purpose

When ANCP is not used, if the BNG needs to manage the line parameters of an AN, the NMS is required. When the AN and the BNG use different NMSs, the line parameters are hard to be managed. Through ANCP, however, the BNG can directly manage such parameters without the NMS.

24.7.2 Principle

The MA5600T supports ANCP for implementing the following functions:

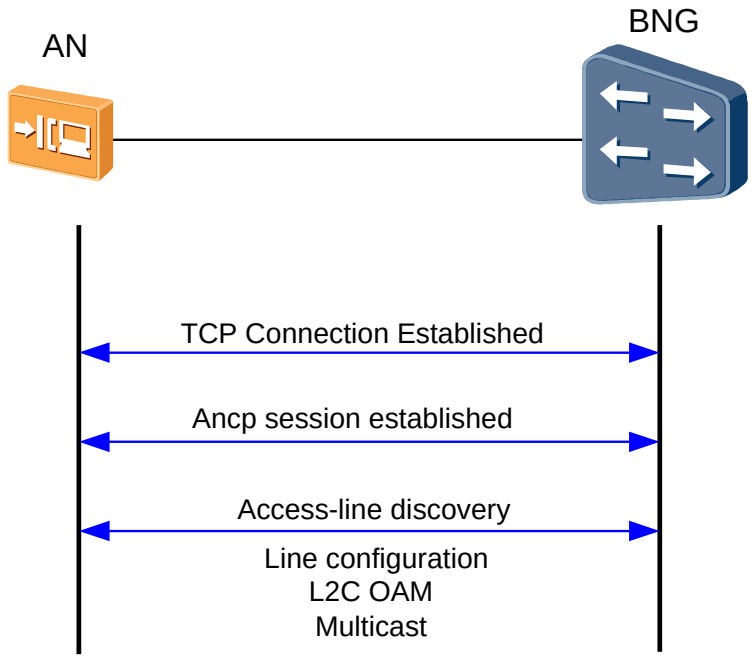
- Line topology discovery
- Line configuration
- L2C OAM
- Multicast and unicast CAC
- Multiple partitioning

Before the above-listed ANCP functions are implemented, an ANCP session needs to be set up between the BNG and the AN.

Setting Up an ANCP Session

Figure 24-9 shows the process of setting up an ANCP session and negotiating the capabilities between the AN and the BNG.

Figure 24-9 Setting up an ANCP session



The process of setting up an ANCP session is as follows:

1. On the AN side, pre-configure the ANCP session IP address and TCP port ID of the BNG and enable ANCP through the CLI. Then, the AN actively sends a request to the BNG for establishing a TCP connection (the BNG is the server and the AN is the client).
2. After the TCP connection is successfully established, adjacency is formed between the AN and the corresponding BNG. After the capabilities are negotiated, the ANCP session is successfully set up. If the local end finds that the remote end does not support a certain capability, the local end disables this capability and negotiates with the remote end again until both ends have negotiated the capabilities supported by both ends.
3. Configure capability parameters. Be default, the AN currently supports the capabilities of line topology discovery, line configuration, and L2C OAM. Multicast and unicast CAC can be added to the capabilities through configuration.

After the adjacency is set up, the ANCP protocol enters the maintaining stage. The AN handshakes with the BNG through the ACK message. The interval is the timeout time contained in the message exchanged during the adjacency setup process. If the AN does not receive the ACK message when the timeout time provided by the BNG expires for three times, the session between the AN and the BNG fails. The AN will then reset the adjacency and initiate a connection again.

Line Topology Discovery

The BNG records the actual parameter information about user ports through line topology discovery and thus implements QoS control.

After a line is activated and the port rate stabilizes, the ANCP module of the AN queries the parameters (such as upstream/downstream activation rate) of the line and sends the port up message and line parameter information to the BNG. After receiving the line information, the BNG saves the information to local and creates mapping to QoS control policies.

After the port is deactivated, the ANCP module sends the port down message to the BNG, as shown in Figure 24-10.

Figure 24-10 Line topology discovery

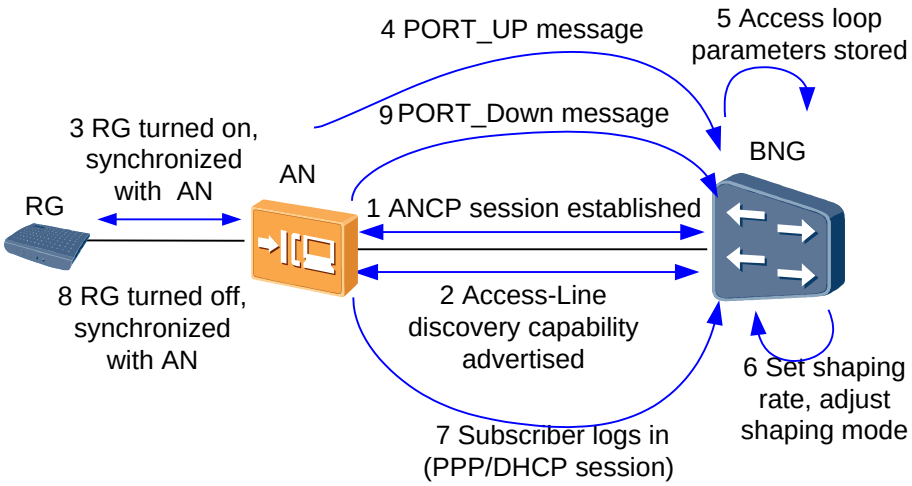


Table 24-5 lists the xDSL line parameters reported by the ANCP module.

Table 24-5 Reported DSL line parameters

No.	Parameter	Meaning
1	DSL Type	Which DSL type is connected (e.g. ADSL, ADSL2, ADSL2+, SHDSL, or VDSL). This parameter defines the transmission system in use.
2	DSL Link State	Line/Port up (Showtime), line/port down (idle or silent)
3	Actual data rate: UPstream and DOWNstream	Actual data rate upstream and downstream of a synchronized DSL link

No.	Parameter	Meaning
4	Attainable Data Rate: UPstream and DOWNstream	Maximum data rate which can be achieved
5	Minimum Data Rate	Minimum data rate desired by the operator in bit/s (up/down)
6	Maximum Data Rate	Maximum data rate desired by the operator in bit/s (up/down)
7	Maximum Interleaving Delay	Maximum one-way interleaving delay
8	Actual Interleaving Delay	Value in milliseconds which corresponds to interleaver setting
9	Minimum-Net-Low-Power-Data-Rate-Upstream	Minimum data rate upstream desired by the operator during the low power state (L1/L2)
10	Minimum-Net-Low-Power-Data-Rate-Downstream	Minimum data rate downstream desired by the operator during the low power state (L1/L2)
11	Access Loop Encapsulation	The link protocol type and the PVC encapsulation of the DSL link

The port up or port down message is uniquely identified by the line ID. The format of the ANCP line ID is configurable. It is recommended to set the format of the ANCP line ID to be consistent with that of DHCP option 82 and PPPoE+ messages. At the same time, the format of the ANCP line ID must be the same as that on the BNG because the BNG creates mapping between user and line according to line ID and user name.

The ANCP module can report the port up or port down message in two modes: based on port or based on service stream.

- In the port-based mode, if SPLABEL (configured by running the **raio-format** command) is not configured in the line ID, the following information is reported when the port is up or down:
  - Message in the default format (VPI=0, VCI=32) if the PVC or CVLAN is not specified
  - Specified value if the PVC or CVLAN specified exists
- If the port-based mode, if SPLABEL (configured by running the **raio-format** command) is configured in the line ID, the following situations occur when the port is up or down:
  - The topology information is not reported if the PVC or CVLAN specified does not exist.
  - The specified value is reported if the PVC or CVLAN specified exists.
- In the service-stream-based mode, when a port goes up or down, messages are reported for all service streams of the port. When the status of a service stream changes, a message is reported for this service stream.

After the ANCP session fails and is re-established, the AN reports the stable line parameters and port status information, such as port UP or port DOWN, to the BNG.

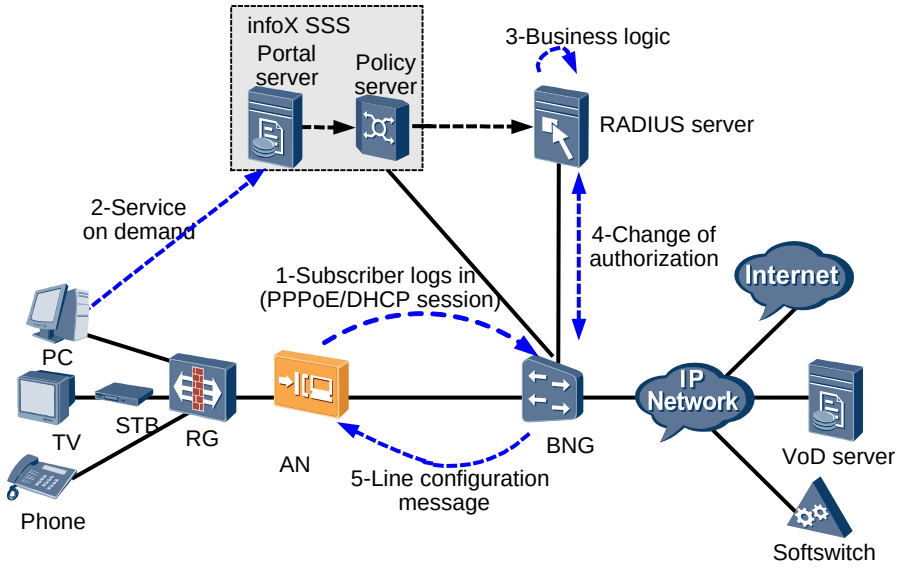
Line Configuration

Most xDSL user parameters are static data. When user service parameters need to be modified, they need to be configured again from the ISP to the access device in an end-to-end manner. Using the ANCP protocol can avoid a complicated exchange process.

The line configuration function is applicable to the self-service customized services. The BNG is required to be able to obtain access line parameters directly from the policy server or RADIUS server and support automatic network update. A precondition is that a copy of line profile parameters of the AN must be saved on the policy server or RADIUS server.

Figure 24-11 shows the process of line parameter modification in a service update.

Figure 24-11 Service update



The process of line parameter modification in a service update scenario is as follows:

1. An ANCP session is established between the AN and the BNG, and a user connects to the BNG.
2. The user orders the required service on the portal server.
3. The portal server and the policy server (through the COPS protocol) or the RADIUS server (through the RADIUS protocol) issues the line ID and the required profile (line profile) name to the BNG. Specific parameters of the profile are already defined on the AN.
4. The BNG issues the received line ID and profile name to the AN through the ANCP protocol.
5. According to the line ID, the AN learns the subrack/slot/port information corresponding to the profile. According to the profile name, the AN knows about the profile to be configured on the port. The AN then uses the new profile to activate the user port to implement the customized service.

NOTE

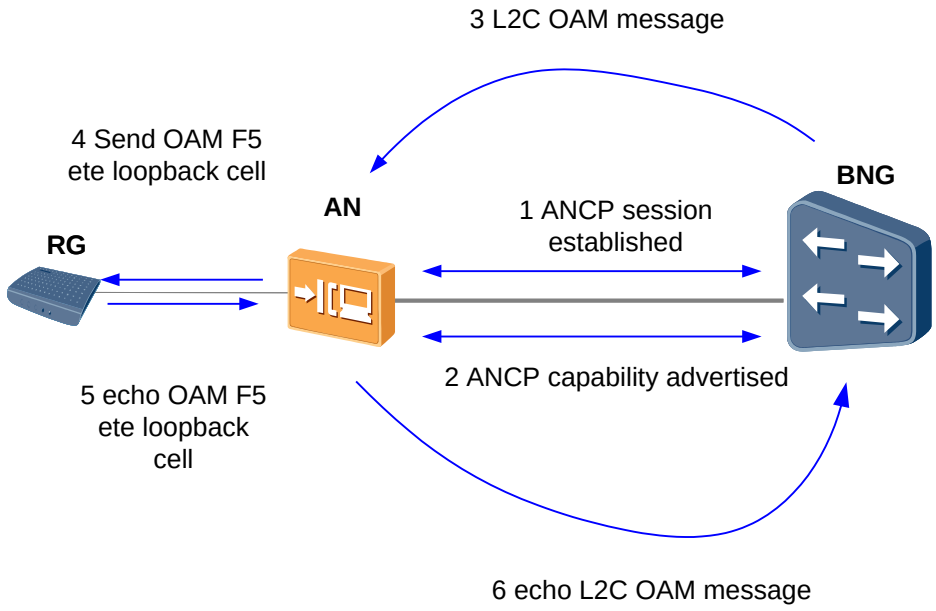
If a line has multiple line profiles, the ANCP supports the configuration of only the profile relevant to the line activation rate.

L2C OAM

The ANCP L2C OAM function can implement the connectivity test between the BNG and the RG.

The implementation of the L2C OAM function mainly involves the RG, AN, and BNG, as shown in Figure 24-12.

Figure 24-12 L2C OAM



L2C OAM in the ADSL and VDSL ATM modes:

- The AN receives the L2C OAM message issued from the BNG and obtains the loopback test port information and the number of loopback cells to be sent (32 by default if a value is not specified). Then, the AN generates loopback cells and sends the cells to the RG.

NOTE

The number of loopback cells ranges from 1 to 32. If the number of loopback cells exceeds 32, the AN discards the excess cells.

- After receiving the loopback response from the RG, the AN obtains the test result. The test result includes the test port information and the loopback result (success or failure). The AN then reports the test result to the BNG through a message.

In the VDSL PTM mode, the AN sends response messages to the BNG according to the port status. If the VDSL port is up, the AN responds with a success message; if the VDSL port is down, the AN responds with a failure message.

Multicast and Unicast CAC

ANCP connection admission control (CAC) and BTV CAC are the same in terms of concept. They are both a protection mechanism for managing the bandwidth of the video programs.

The ANCP multicast CAC and unicast CAC mainly involve three devices: AN, BNG, and resource admission control subsystem (RACS). The AN and the BNG are connected through ANCP, and the BNG and the RACS are connected through COPS interfaces.

Video bandwidth:

Total video bandwidth of AN users = Multicast program bandwidth of the users + Unicast VoD bandwidth of the users

- User multicast bandwidth CAC is implemented on the AN.
- User unicast VoD bandwidth CAC is implemented on the RACS (the policy server on the Figure 24-13).

Video bandwidth waterline mechanism:

The AN introduces a video bandwidth waterline mechanism to dynamically adjust the video bandwidth between the AN and the RACS.

In this mechanism, the total video bandwidth of users is compared to a container. Through the video bandwidth waterline mechanism, the waterline in the container can be upshifted/downshifted to dynamically adjust the bandwidth resources for multicast programs and the bandwidth resources for unicast VoD programs. When the bandwidth resources for multicast programs are insufficient, the bandwidth resources for unicast VoD programs can be requested for multicast programs. The same is true for unicast VoD programs. This mechanism enhances user experience with program demanding.

Video bandwidth waterline data update:

When the AN or the RACS applies to each other for bandwidth, the applicant starts a timer (set to 500 ms). If the applicant does not receive a response after the timer times out, the AN deletes the IGMP join message of the corresponding user from the buffer and returns a program demand failure to the user. At the same time, the AN or the RACS actively queries each other about the video bandwidth information and updates their own video bandwidth waterline data according to the waterline data of the peer. If the query about the video bandwidth information fails, the AN or the RACS does not update the video bandwidth waterline data.

ANCP feature and BTV feature:

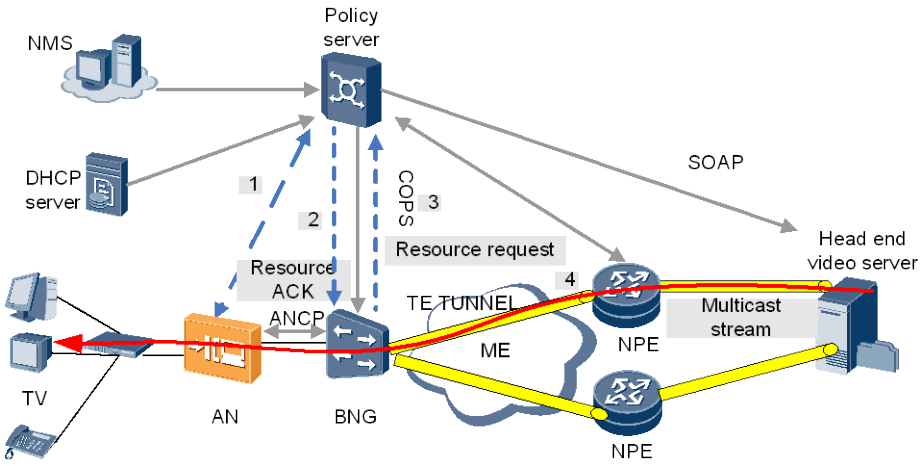
The relationship between the ANCP multicast CAC feature and the BTV multicast CAC feature is as follows:

- When BTV multicast CAC is not enabled, the AN cannot implement multicast CAC regardless of whether ANCP multicast CAC is enabled or not.
- The video bandwidth between the AN and the RACS can be dynamically adjusted only when BTV multicast CAC and ANCP multicast CAC are enabled.

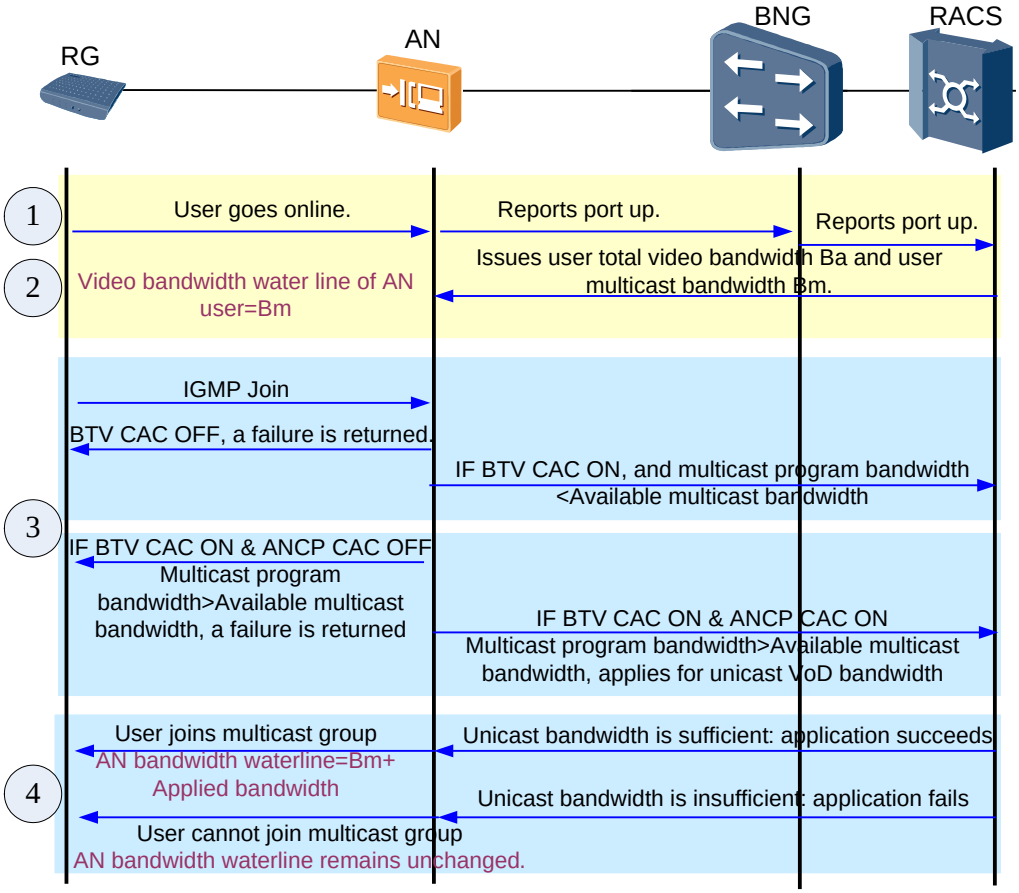
Multicast CAC

Figure 24-13 shows the application scenario of multicast CAC.

Figure 24-13 Application scenario of multicast CAC



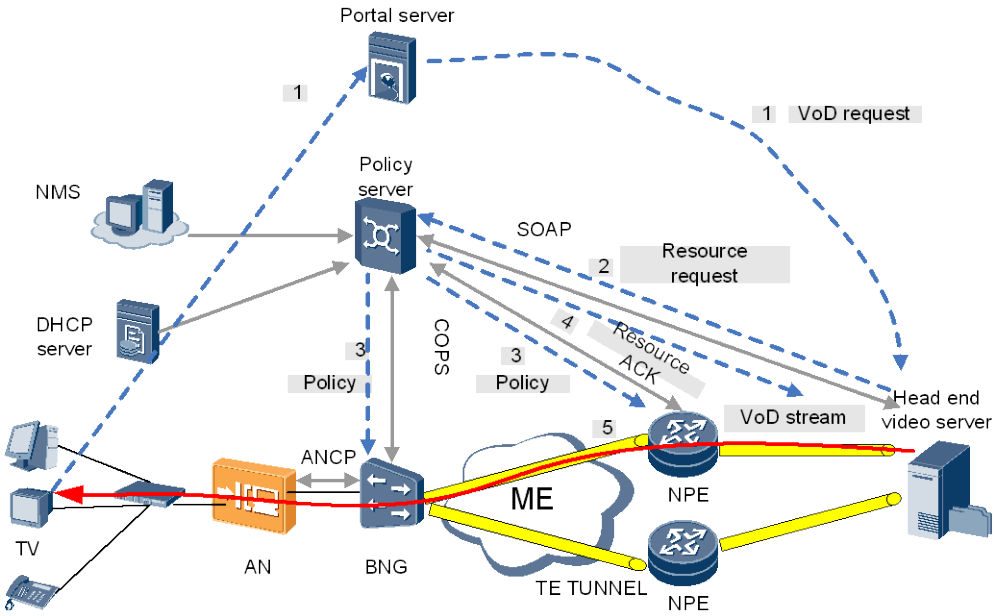
The message exchange process of multicast CAC is as follows:



Unicast CAC

Figure 24-14 shows the application scenario of unicast CAC.

Figure 24-14 Application scenario of unicast CAC



After detecting that a user goes online, the RACS actively configures the user video bandwidth information for the AN and updates the user video bandwidth waterline on the AN according to the user video bandwidth information on the RACS.

The user video bandwidth waterline on the RACS equals the user unicast VoD bandwidth. If the total video bandwidth equals the unicast VoD bandwidth, the RACS manages all the video bandwidth for the user. In this case, the multicast bandwidth of the user is 0.

The message exchange process of unicast CAC is as follows:

1. When the user demands a unicast program through the portal interface, the demand information is transmitted to the VoD server through the data channel.
2. The VoD server requests unicast CAC from the RACS. The RACS compares the unicast VoD bandwidth requested by the user with the available unicast VoD bandwidth of this user.
   - If the unicast VoD bandwidth requested by the user is smaller than the available unicast VoD bandwidth, the user is allowed to demand this unicast VoD program and the available unicast VoD bandwidth is updated on the RACS. After the update, the available unicast VoD bandwidth of the user = Pre-update unicast VoD bandwidth of the user - Unicast VoD bandwidth requested by the user.
   - If the unicast VoD bandwidth requested by the user exceeds the available unicast VoD bandwidth, the RACS needs to apply to the AN for multicast bandwidth resources through the extended ANCP message.
3. The ANCP module of the AN will receive the bandwidth application message. If the multicast CAC of the ANCP module is disabled, the AN responds with a request failure message. If multicast CAC is enabled on the ANCP module, the ANCP module checks whether the available multicast bandwidth of the user is sufficient for the bandwidth requested by the unicast user.
   - If the remaining available multicast bandwidth is sufficient, the ANCP module grants the multicast bandwidth to the unicast user and then sends an ANCP message to the RACS to notify the success, and at the same time updates the AN video

bandwidth waterline. The bandwidth requested by the RACS this time needs to be deducted from the video bandwidth waterline and also needs to be deducted from the remaining available multicast bandwidth of the user.

- If the remaining available multicast bandwidth is insufficient, applying for the multicast bandwidth fails. In this case, the ANCP module sends an ANCP message to the RACS to notify the failure.

4. The RACS processes the bandwidth application results accordingly. If the application is successful, the RACS updates the unicast bandwidth on the RACS (the original unicast bandwidth + the successfully requested bandwidth this time), and returns the unicast CAC result to the VoD server.
5. The VoD server processes the unicast CAC result. If the VoD server processes the unicast CAC result successfully, demanding the unicast program continues. Otherwise, demanding the unicast program is stopped.

NOTE

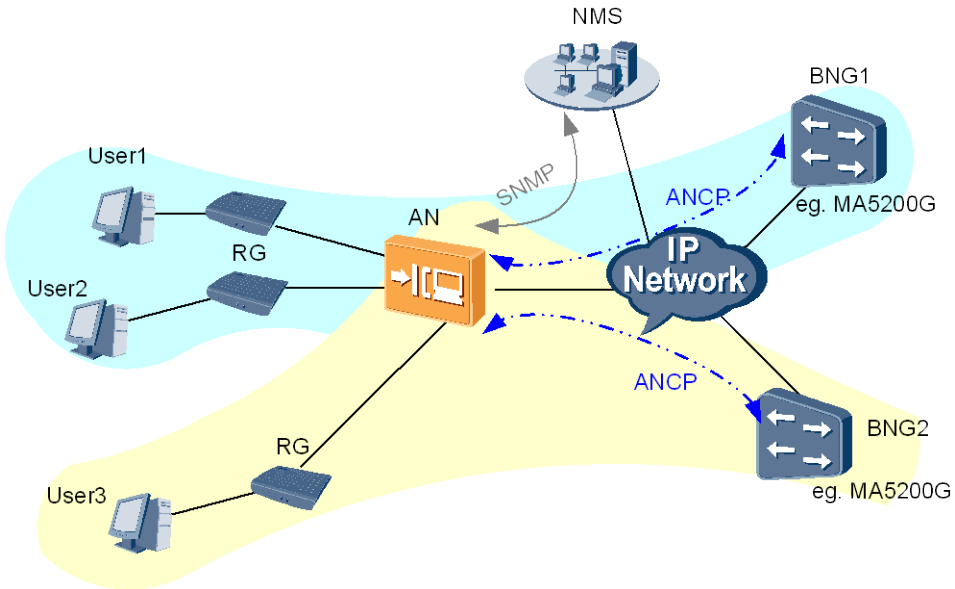
If the user stops demanding VoD programs, the VoD server sends the RACS a message indicating that demanding VoD programs has stopped. In this case, the RACS updates the unicast bandwidth on the RACS (the unicast bandwidth + the bandwidth of this VoD program).

Multiple Partitioning

The ANCP multiple partitioning function enables different xDSL ports to the grouped in different ANCP partitions. Different partitions are managed by different BNGs in order to support wholesale service. Generally, each BNG is managed by a respective ISP. One xDSL port can belong to only one partition.

As shown in Figure 24-15, user 1 and user 2 are managed by BNG 1, and user 3 by BNG 2. In other words, user 1 and user 2 belong to the same partition. and user 3 belong to a different partition.

Figure 24-15 Network topology of ANCP multiple partitioning



24.7.3 Configuring ANCP

Access Node Control Protocol (ANCP) is used to implement the functions such as topology discovery, line configuration, and L2C OAM on the user ports. The MA5600T/MA5603T/MA5608T establishes an ANCP session according to the GSMP communication IP address configured in the network access server (NAS).

Prerequisites

- The system must work in the normal state.
- The system must be connected to the network access server in the normal state.

Context

- The MA5600T/MA5603T/MA5608T and the NAS use the TCP connection to carry an ANCP session. Therefore, before creating the ANCP session, you must create a TCP connection between the MA5600T/MA5603T/MA5608T and the NAS. The NAS functions as the server of the TCP connection, and the MA5600T/MA5603T/MA5608T functions as the client of the TCP connection.
- After the TCP connection is created successfully between the MA5600T/MA5603T/MA5608T and the NAS, an ANCP session is created between the MA5600T/MA5603T/MA5608T and the NAS. After the ANCP session is created successfully, the MA5600T/MA5603T/MA5608T and the NAS need to use the ANCP ACK packets for heartbeat detection to maintain the ANCP session.
- The default values of the ANCP parameters are as follows:
  - GSMP address for an ANCP session: 0.0.0.0
  - ANCP session capability set: topology-discovery, line-config, and oam
  - ANCP packet sending priority: 6
  - GSMP TCP communication port number on the NAS side in an ANCP session: 6068
  - Interval for sending packets during the initial stage of an ANCP session: 10 (unit: 0.1s)
  - Interval for sending packets during the ANCP session stage: 100 (unit: 0.1s)

Procedure

Run the **ancp partition enable** command to enable the ANCP partition function.

By default, the ANCP partition function is disabled.

Step 1 Run the **ancp port** command to enable the ANCP function of a port.

The ANCP function takes effect only when the ANCP function in the ANCP session mode and ANCP session function of a port are enabled.

Step 2 (Optional) Run the **ancp version** command to configure the ANCP version.

- The configured ANCP version must be the same as that on the NAS.
- By default, the ANCP version is draft-01.

Step 3 (Optional) Run the **raio-format ancp aggregation-circuit-id** command to customize the character string format for Access-Aggregation-Circuit-ID-ASCII TLV of the ANCP message.

NOTE
The system supports to customize the character string format for Access-Aggregation-Circuit-ID-ASCII TLV of the ANCP message only when the ANCP version is RFC6320.

Step 4 Run the **ancp session** command to create the ANCP session.

Step 5 (Optional) Run the **ancp partition** command to configure the ID of the partition associated with an ANCP session.

Step 6 Run the **ancp ip** command to configure the GSMP communication IP address for the ANCP session.

- The IP address configured here must be the same as the GSMP communication IP address configured on the NAS, but it should to not be the same as the default IP address, multicast IP address, or broadcast IP address.
- When an ANCP session is enabled, the GSMP communication IP address cannot be configured.

Step 7 (Optional) Run the **ancp capability** command to configure the capability set of the ANCP session.

- Supports topology discovery. When you select **topology-discovery** parameter, the MA5600T/MA5603T/MA5608T automatically reports the line parameters to the NAS.
- Supports line configuration. When you select **line-config** parameter, the MA5600T/MA5603T/MA5608T responds to the line configuration that is sent by the NAS.
- Supports the OAM. When you select **oam** parameter, the MA5600T/MA5603T/MA5608T responds to the line testing information that is sent by the NAS.
- Supports the preceding three types of capability.
- The default value is all, that is, the three capabilities (topology discovery, line configuration, and L2C OAM) are supported.

NOTE
When the ANCP version is RFC6320, the system only supports **topology discovery**.

Step 8 (Optional) Run the **ancp ancp-8021p** command to set the priority for sending ANCP packets.

- You can set the priority according to the actual requirements and network conditions, the higher the priority, the higher the reliability.
- After an ANCP session is enabled, the priority for sending the ANCP packet of the ANCP session cannot be configured.

Step 9 (Optional) Run the **ancp nas-tcp-port** command to set the GSMP TCP communication port number for the ANCP session on the NAS.

- By default, the GSMP TCP communication port number is 6068.
- The GSMP TCP communication port number on the MA5600T/MA5603T/MA5608T must be the same as that on the NAS.
- Run the **ancp port begin** command to set the start port ID of the ANCP session. Make sure that the start port ID of the ANCP session is the same as the start ID of the ports on the service board.

Step 10 (Optional) Run the **ancp init-interval** command to set the interval for sending packets during the establishment of the ANCP session.

- By default, the general query interval is 125s.

- After an ANCP session is enabled, the priority for sending the ANCP packet of the ANCP session cannot be configured.

Step 11 (Optional) Run the **ancp keep-alive** command to set the interval for sending packets during the ACNP session so that the handshake messages can be sent to the peer end at the preset interval.

- By default, the interval is 10s.
- After an ANCP session is enabled, the priority for sending the ANCP packet of the ANCP session cannot be configured.

Step 12 (Optional) Run the **ancp bandwidthCAC** command to enable the ANCP multicast CAC. After the ANCP multicast CAC is enabled, if the bandwidth of the demanded multicast program is larger than the available multicast bandwidth of the user, the user can apply for the bandwidth resource of the unicast VOD program.

- After an ANCP session is enabled, its ANCP multicast CAC function cannot be enabled or disabled.
- The ANCP multicast CAC function of only one session can be enabled at a time.
- After the ANCP multicast CAC function is enabled, if the **ancp disable** command is executed, the ANCP will be disabled. The system still performs CAC using the bandwidth issued by the ANCP CAC and the original BTV CAC does not take effect. In this case, the normal BTV CAC takes effect only when the ANCP CAC function of the ANCP session is disabled by running the **ancp bandwidthCAC disable** command.

Step 13 Run the **ancp enable** command to enable the ANCP session.

- By default, the ANCP session is disabled.
- Before an ANCP session is enabled, related parameters can be modified. After an ANCP session is enabled, related parameters cannot be modified.

Step 14 Run the **quit** command to quit the ANCP mode.

Step 15 Run the **display ancp session** command to query the information about the ANCP session.

----End

Example

Consider configuring the ANCP topology discovery function of port 0/3/1 as an example. Configure the partition ID of the ANCP session to 1, ANCP version to rfc6320, the character string format for Access-Aggregation-Circuit-ID-ASCII TLV of the ANCP message is innervlanid and outervlanid, start port ID to 1, GSMP communication address of the ANCP session to 10.10.10.10, packet sending interval at the ANCP session creation phase to 2s, ANCP session capability set to topology-discovery, ANCP packet sending priority to 7, GSMP TCP communication port ID at the NSA side in the ANCP session to 6000, and packet sending interval at the ANCP session phase to 7s.

```
huawei(config)#ancp partition enable
huawei(config)#ancp port 0/3/1 partition 1
huawei(config)#ancp version rfc6320.
huawei(config)#raio-format ancp aggregation-circuit-id innervlanid.outervlanid
huawei(config)#ancp port begin 1
huawei(config)#ancp session 1
huawei(config-session-1)#ancp partition 1
huawei(config-session-1)#ancp ip 10.10.10.10
```

```
huawei(config-session-1)#ancp capability topology-discovery
huawei(config-session-1)#ancp ancp-8021p 7
huawei(config-session-1)#ancp nas-tcp-port 6000
huawei(config-session-1)#ancp init-interval 20
huawei(config-session-1)#ancp keep-alive 70
huawei(config-session-1)#ancp bandwidthCAC enable
huawei(config-session-1)#ancp enable
huawei(config-session-1)#quit
huawei(config)#display ancp session 1

  Session config status              : Enable
  Session running status             : Before syn phase
  Session diagnostic status          : -
  GSMP version                      : 3
  GSMP sub version                  : 1
  AN name                           : -
  NAS name                          : -
  NAS IP                            : 10.10.10.10
  Local IP                          : -
  AN instance                       : -
  NAS instance                      : -
  Config capabilities               : TopologyDiscovery
  Negotiate capabilities            : -
  NAS TCP port                      : 6000
  Startup time(0.01s)               : -
  Discontinuity time(0.01s)         : -
  Init interval(0.1s)               : 20
  Keepalive interval(0.1s)          : 70
  PartitionID                       : 1
  Bandwidth CAC status              : Enable
  Line config roll default          : Disable
  OAM threshold(0.01)               : 100
  Topology report shaper interval(0.1s) : 10
  S-VLAN                            : -
  S-VLAN priority                   : 7
  C-VLAN                            : -
  C-VLAN priority                   : -
  Session down send trap status      : Disable
  Session up send trap status        : Disable
```

24.8 Remote Connection Security

This topic provides an introduction to the remote connection security, and describes the working principle of this sub feature.

24.8.1 Introduction

Definition

With the remote connection security feature, the IP firewall, or the service port of the system is disabled to prevent the device from being attacked by illegal users or illegal operations.

Purpose

IP firewall or disabling the service port can prevent the device from being attacked by malicious users to ensure the security of the device.

24.8.2 Principle

With the IP firewall function, only the operators from valid IP address segments are allowed to log in to the device through valid access protocols, and the operators from invalid IP address segments or through invalid access protocols are not allowed to log in to the device.

With the function of disabling the system service, the default service monitoring port of the system can be disabled to prevent the port from malicious scanning or attack.

24.9 Log Management

This topic covers the overview and working principle of log management.

24.9.1 Introduction

Definition

Logs can be classified into security event logs and operation logs.

- A security event log is a log recorded by the system after a security event occurs. Currently, three types of security events are supported, that is, online/offline event of maintenance users, user lockout event, and auto-backup success event.
- An operation log is a log about the user operation recorded by the system. It records user login and logout information and other operations performed on the system.

Generally, logs are queried through the CLI, syslog, or backup log file during troubleshooting.

Operation logs and security event logs are reported to the NMS.

Purpose

Logs recorded help users obtain the overall system maintenance information for timely troubleshooting.

24.9.2 Principle

Operation Log

The system records commands of successfully issued configurations from the CLI or SNMP interface, that is, operation logs. Operation logs record both successful and failed operations. In logs of failed operations, the operation results can also be recorded.

By default, the system supports a maximum of 5000 operation logs, which are saved in the order of time and are overwritten cyclically. After system restart, logs recorded are not lost.

Security Event Log

Events are reminders to the user during the system running. The event attributes include the event ID, event name, event type, event class, event level, and the default event level, where the event level can be customized.

When the level of a security event is changed, whether the event is recorded may be changed. A security event is recorded in the log only when its level is minor or higher.

Log Server

Logs can be reported to the log server using syslog in real time. Also, logs can be transmitted to the file server through TFTP/FTP/SFTP at a specified time or when the specified capacity is reached after the automatic uploading conditions are configured. Integrity of logs must be ensured.

NOTE
SFTP is recommended.

NMS Log Management

NMS log management involves management of NMS security logs, NMS operation logs, and NE security logs. By querying and saving logs periodically, network management engineers can detect unauthorized logins or operations and analyze faults in time. Through the logs, the information about the client from which the NMS user logs in to the NMS server and the operations performed after login can be obtained. Also, log data can be dumped or printed.

24.10 Version and Data Management

This topic provides an introduction to the version and data management feature, and describes the working principle of this sub feature.

24.10.1 Introduction

Definition

Version and data management includes patch management, rollback function, configuration data management, and version upgrade.

Purpose

This sub feature facilitates carriers in version upgrade and maintenance.

Benefits

Benefits to carriers: The carriers' operating expenditure (OPEX) is saved considerably, and the customer satisfaction is increased.

24.10.2 Principle

Patch Management

The flash memory (storage medium in the system) has a patch area to store the loaded patches.

A patch can be a hot patch or cold patch. The system needs to be restarted for a cold patch to take effect or stop functioning. Nevertheless, in the case of a hot patch, the system need not be restarted for the same purpose. A hot patch supports the rollback function; therefore, the hot patch can be rolled back to the status before the latest hot patch is loaded.

In addition, a patch can be activated, deactivated, run, or deleted. The loaded patch is deactivated by default; therefore, to make the loaded patch take effect, activate it. To make the patch take effect after the system restart, activate and run the patch before the system is restarted.

The system supports the following four types of patches:

- HP refers to the host hot patch. It takes effect after being loaded and then activated. For a user, this type of patches, after being loaded, is displayed as HPXXX.
- SPH is the set of HP patches. It takes effect after being loaded and activated. For a user, this type of patches, after being loaded, is displayed as SPHXXX, without displaying the status of HP patches
- CP refers to the host cold patch. It takes effect after it is loaded and the system is restarted. For a user, this type of patches, after being loaded, is displayed as CPXXX.
- SPC is the set of CP patches. It takes effect after being loaded and activated. For a user, this type of patches, after being loaded, is displayed as SPCXXX, without displaying the status of CP patches.

Rollback Function

The flash memory of the control board is divided into two same storage areas (namely, active storage area and standby storage area) to store the program, database, and extended BIOS. The storage area that is operating currently is the active storage area. When the program, database, and extended BIOS are upgraded, the new program, database, and extended BIOS are loaded to the standby storage area. After the system is restarted, the system automatically loads the new program, database, and extended BIOS. The rollback function is implemented based on two sets of program, database, and extended BIOS in both the active and standby storage areas.

By default, after upgrade, the system saves the pre-upgrade host program and database for 48 hours. 48 hours later, the system automatically cancels the rollback function. That is, 48 hours later, the system duplicates the program, database, and extended BIOS in the operating area to the standby storage area. In this way, the versions in both the active and standby storage areas are the same. You can set the time for canceling the rollback function to 5 minutes to 30 days.

The system supports automatic rollback and manual rollback. After version upgrade, if the system fails to start up, the system is automatically rolled back to the version before upgrade. After version upgrade, if the system becomes abnormal during the running and cannot recover, you can run the rollback command to roll back the system to the version before upgrade.

Configuration Data Management

- Saving the configuration data manually: The current configuration data can be saved manually through the commands. If the configuration data is not saved before the system is reset or restarted, it will be lost after the system reset or restart. Therefore, manually save the configuration data once before the system is reset or restarted.
- Saving the configuration data after any changes to the configuration data: After the configuration data is changed, the system will save the changed configuration data automatically at a preset interval. This interval is user-defined and ranges from 10 minutes to 10080 minutes (default value: 30minutes).
- Saving the configuration data at a preset time or interval: In the system, the configuration data can be saved automatically at a preset time or interval. This time or interval is user-defined. For example, the time or interval can be set to 23:00 or two hours respectively. In this case, the configuration data is saved at an interval of two hours or at 23:00.

The data erasure operation can be performed to restore the configuration data of the device to the default settings.

The system also supports backing up the current configuration data manually or at a preset time to a specified file server.

Version Upgrade

Software version in the system can be upgraded through the CLI or the NMS by using FTP/TFTP/XMODEM/STFP.

NOTE
SFTP is recommended.

24.11 LLDP

Link Layer Discovery Protocol (LLDP) is a standard link layer discovery mode defined in IEEE 802.1ab. LLDP deployed on multi-vendor devices that are running in different networks and managed by the network management system (NMS) enables adjacent network devices to exchange device information with each other. By accessing the device information, the NMS obtains detailed information, such as topology of the whole network and physical connections between devices, in real time. As such, LLDP helps users monitor network status and locate network faults.

24.11.1 Introduction

This topic describes the background and application values of LLDP.

As large-scale networking requirements keep emerging, a network is involving an increasingly large number of device types, each device having its own complicated configurations. In addition, higher requirements are posed on the NMS capabilities. For example, the NMS is required to automatically learn the topological status of connected devices, and to detect configuration conflicts between devices.

Most NMSs use the automated discovery function to trace network topology changes, but in this way the NMSs can trace the topology to as deep as the network layer only. As such, operators can learn only the basic events, such as device addition or deletion, on a network, but cannot identify the specific location of a device. For example, operators cannot identify

the ports through which a device connects to other devices. Though some vendors provide proprietary protocols for discovering adjacency between devices, there are still not sufficient means for generating the topology of an entire network.

LLDP well addresses the issues above. As a Layer 2 discovery protocol defined in IEEE 802.1ab, LLDP provides a standard link layer discovery method. Using this method, information such as the capabilities, management address, device ID, and port ID of a local device can be encapsulated in LLDP frames and transmitted to adjacent nodes. After receiving such information, the adjacent nodes save the information in a standard management information base (MIB), which can be queried by the NMS for determining the connectivity of links.

LLDP can precisely discover the ports of network devices and the interconnection between devices. By obtaining and integrating the LLDP local device information and neighbor device information on each network element (NE), the NMS can generate a clear topology of the entire network, along with detailed information such as the physical connections between devices. Such information helps network operators monitor network status in real time and quickly locate network faults, effectively improving network security and stability.

24.11.2 Basic Concepts

This topic describes basic concepts related to the LLDP feature to help you better understand the working principles of LLDP.

LLDP MIB

MIB is a fundamental element for LLDP implementation. As defined in the LLDP protocol, each port on a device has four MIBs. Among them, the two most important MIBs are the LLDP local system MIB and the LLDP remote system MIB, which store the status information of the local device and the adjacent node, respectively. The status information includes device ID, port ID, system name, system description, port description, device capabilities, and network management address.

LLDP Agent

The LLDP protocol defines an LLDP agent for each port on a device for managing the LLDP operations regarding the device. The LLDP agent:

- Maintains the LLDP local system MIB information.
- Sends LLDP frames to the adjacent nodes, advertising the local system status information to the adjacent nodes.
- Recognizes and parses the LLDP frames sent from the adjacent nodes, and maintains the LLDP remote system MIB information.
- Sends LLDP alarms to the NMS when the LLDP local system MIB or LLDP remote system MIB information changes.

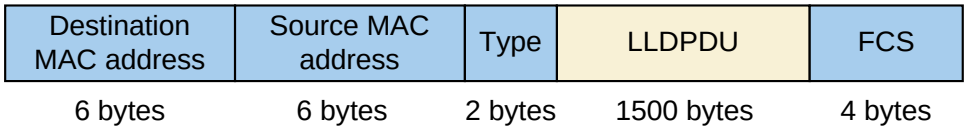
Compared with the IEEE 802.1ab-2009 specifications, Huawei's LLDP implementation supports one single port associated with one LLDP agent, but does not support one single port associated with multiple LLDP agents.

LLDP Frame

Ethernet frames in which LLDP data units (LLDPDUs) are encapsulated are called LLDP frames. The encapsulation has two formats: Ethernet II format and Subnetwork Access Protocol (SNAP) format. The MA5600T/MA5603T/MA5608T supports the Ethernet II

encapsulation format. The following figure shows the structure of the LLDP frame encapsulated in this format.

Figure 24-16 LLDP frame structure



The following table explains the fields of the LLDP frame.

Field	Description
Destination MAC address	Set to a fixed multicast MAC address 0x0180-C200-000E, which is specified by IEEE 802.1ab-2005. There is usually no two-port MAC relay (TPMR) between the MA5600T/MA5603T/MA5608T and the connected devices. Therefore, for the purpose of maximal compatibility (that is, considering all devices as common devices), the MA5600T/MA5603T/MA5608T uses 01-80-C2-00-00-0E as the destination MAC address. For the requirements of different types of devices on the reception and transmission of the destination MAC address, see IEEE 802.1ab-2009.
Source MAC address	The device MAC address serves as the source MAC address.
Type	Indicates the frame type, which is a fixed value 0x88CC.
LLDPDU	Indicates the LLDP data unit. LLDP information exchange is implemented through LLDPDUs.
FCS	Indicates the frame check sequence.

LLDPDU

LLDPDU is a data unit that is encapsulated in an LLDP frame. A device encapsulates the local information in the Type-Length-Value (TLV) format. Several such TLVs constitute an LLDPDU and are transmitted in an LLDP frame. Users can formulate LLDPDUs using different combinations of TLVs based on their requirements. Using these TLVs, the device advertises its status and learns the status of its adjacent nodes.

Figure 24-17 LLDPDU structure

The LLDP protocol stipulates four mandatory TLVs for an LLDPDU. Each LLDPDU must start with the Chassis ID TLV, Port ID TLV, and Time to Live TLV, and end with the End of LLDPDU TLV. Other TLVs are optional.

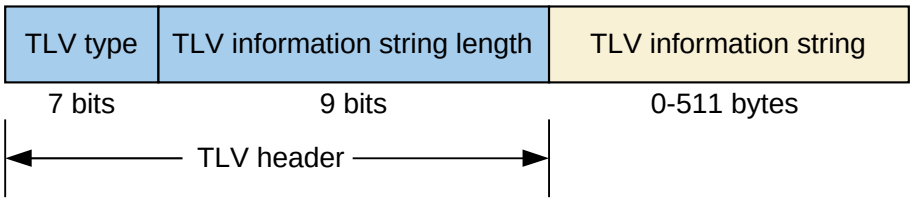
There are two types of LLDPDUs:

- Normal LLDPDU: contains the local device information and the adjacent node information.
- Shutdown LLDPDU: sent to disable the LLDP capability of a port. The shutdown LLDPDU is used to instruct the peer device to quickly clear the adjacent node information. The shutdown LLDPDU does not have optional TLVs, and the TTL TLV value is 0.

TLV

Indicating the type, length, and value of an object, TLV is the basic unit that forms an LLDPDU. Each TLV represents a type of device information, such as device ID, port ID, and management address, which correspond to fixed TLVs Chassis ID TLV, Port ID TLV, and Management Address TLV respectively.

Figure 24-18 TLV structure



LLDP supports two TLV formats: basic TLV and organizationally specific TLV. Organizationally specific TLVs include TLVs defined by 802.1 and 802.3, and may include more organizationally defined TLVs. The MA5600T/MA5603T/MA5608T supports only the basic TLV.

Table 24-6 TLV list

TLV Name	TLV Type	Description
End of LLDPDU TLV	0	Indicates the end of an LLDPDU.
Chassis ID TLV	1	Indicates the bridge MAC address of the device.
Port ID TLV	2	Indicates the name of the LLDPDU sending port. The value of this TLV uses the value of the ifName leaf of the IF-MIB.
Time To Live TLV	3	Indicates the time during which the local device information is valid on the adjacent node.
Port Description TLV	4	Indicates the port description. The value of this TLV uses the value of the ifDescr leaf in the iftable table of the IF-MIB.
System Name TLV	5	Indicates the device name, configurable using the **sysname** command.

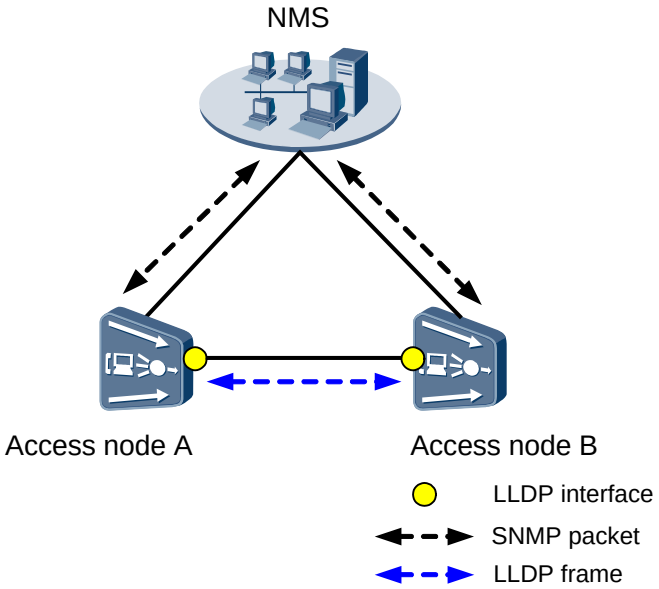
TLV Name	TLV Type	Description
System Description TLV	6	Indicates the device description, configurable using the **system sys-info description** command.
System Capabilities TLV	7	Indicates the functions supported by the device and the functions that are enabled on the device.
Management Address TLV	8	Indicates the management address.
Reserved	9-126	Reserved for special application.
Organizationally Specific TLV	127	Each organization is represented by an organizationally unique identifier (OUI). For detailed description of this TLV and the OUI field, see the LLDP protocol.

24.11.3 Principles

Based on an LLDP basic network diagram, this topic describes how the NMS obtains device information and learns the network topology using the LLDP protocol.

LLDP Working Process

Figure 24-19 LLDP basic network diagram



As shown in the preceding figure, access node A and access node B support LLDP. The NMS collects device information in the following process:

1. Access node A encapsulates its local status information in an LLDP frame and sends the frame to its adjacent device, access node B.
2. Access node B receives and parses the LLDP frame, and stores the information about access node A in its LLDP remote system MIB. The information can be accessed by the NMS when the NMS extracts the topology information.
3. Similarly, access node B encapsulates its local status information in an LLDP frame and sends the frame to access node A.
4. Access node A receives and parses the LLDP frame, and stores the information about access node B in its LLDP remote system MIB. The information can be accessed by the NMS when the NMS extracts the topology information.
5. The NMS exchanges SNMP messages with access node A and access node B, and extracts the local device information and adjacent device information from their MIBs. By integrating and analyzing the information, the NMS discovers the topology of the whole network.

In the working process described above:

- To generate the topology of the entire network on the NMS, all devices managed by the NMS must support LLDP. LLDP is disabled on the MA5600T/MA5603T/MA5608T by default.
- Each device can discover the information about only the device that is directly connected to this device. Therefore, to generate a whole-network topology, the NMS needs to collect the local device information and adjacent device information reported by all devices on the network.

LLDP Port Working Modes

An LLDP port supports the following working modes:

- Tx/Rx: The port transmits and receives LLDP frames. This is the default mode on the MA5600T/MA5603T/MA5608T.
- Tx: The port transmits but does not receive LLDP frames.
- Rx: The port receives but does not transmit LLDP frames.
- Disable: The port does not transmit or receive LLDP frames.

When the LLDP working mode of a port changes, the port initializes the protocol state machine. If the LLDP working mode of a port changes frequently, the port has to constantly perform re-initialization. To avoid this situation, the MA5600T/MA5603T/MA5608T supports configuration of a port initialization delay. A port performs re-initialization only when the delay expires after the port working mode changes.

LLDP Frame Transmission Mechanism

An LLDP-enabled device periodically transmits LLDP frames to its adjacent nodes. To make sure that it is detected by other devices as quickly as possible, the MA5600T/MA5603T/MA5608T supports a fast transmission mechanism. The MA5600T/MA5603T/MA5608T immediately transmits an LLDP frame in any of the following three conditions:

- The MA5600T/MA5603T/MA5608T discovers a new neighbor. In other words, the MA5600T/MA5603T/MA5608T receives a new LLDP frame and finds that the transmitting device information carried in the frame is not stored on the MA5600T/MA5603T/MA5608T yet.

- The LLDP status of the MA5600T/MA5603T/MA5608T changes from disabled to enabled.
- The port status of the MA5600T/MA5603T/MA5608T changes from Down to Up.

LLDP Frame Reception Mechanism

When receiving an LLDP frame, the device checks the validity of the frame and the TLV information carried in the frame. After the frame passes the validity check, the device stores the neighbor's information. According to the Time To Live TLV value carried in the LLDPDU, the device sets the aging time of the neighbor's information on this device. When receiving an LLDPDU whose Time To Live TLV value is 0, the device ages the information of the corresponding neighbor.

24.11.4 Network Application

The LLDP topology discovery function is usually used with the function of remote software commissioning (through GE upstream channel) to implement the plug-and-play (PnP) solution.

A general PnP solution is implemented in three steps:

1. After a device (connected to the upstream through GE) is powered on, the device automatically obtains its IP address and management VLAN parameters. Such information is used for setting up the management channel for the device to enable remote management.
2. Service configurations of the device are automatically issued to the device; automatic upgrade of the device is pre-deployed.
3. The network topology is discovered on the NMS.

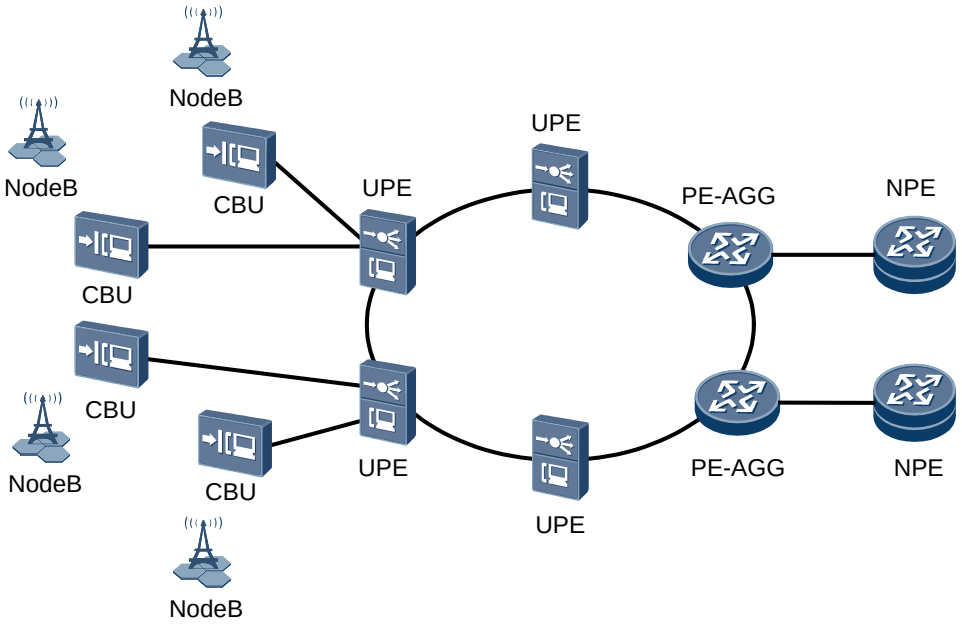
The first step is implemented using the function of remote software commissioning (through GE upstream channel). With this function, operators need not visit the sites to commission software, thereby reducing the costs for deploying a large number of devices. Steps 2 and 3 require coordination of the NMS. The LLDP function works with the NMS to implement step 3.

According to how the access device is connected to upstream devices, topology discovery of the access device includes the following scenarios.

Point-to-Point Connection

As shown in Figure 24-20, the access node MxU works as the CBU, and an OLT or a switch works as the UPE. The MxU connects to the upstream device through FE/GE. The IP address is obtained using DHCP, and the device location is identified by DHCP option 82. Specifically, DHCP relay or DHCP proxy is enabled on the upper-layer device, and the device reports the MxU location information to the NMS through DHCP option 82. The topology of this type of network can be discovered using LLDP.

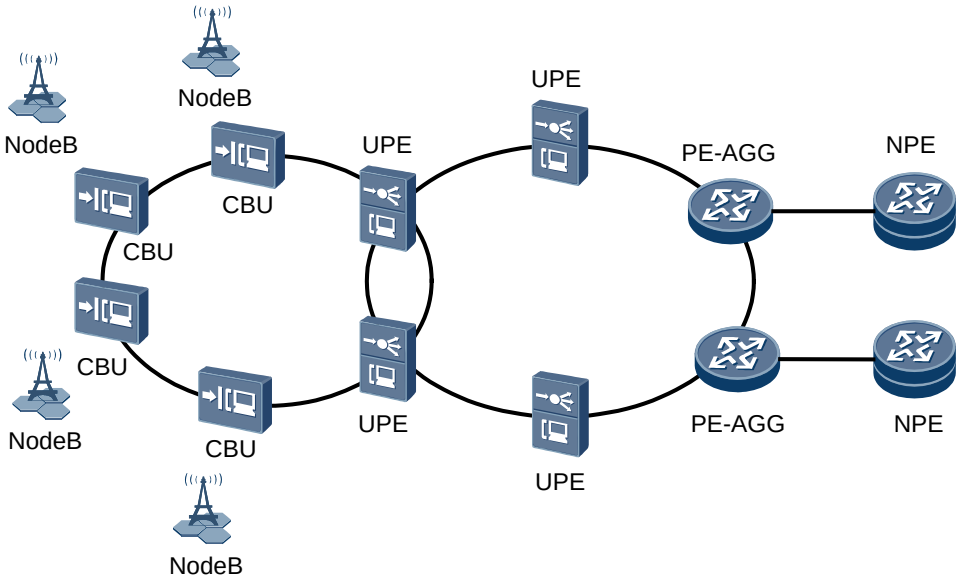
Figure 24-20 Network of point-to-point connection



Ring or Chain Network Topology

For a network topology shown in Figure 24-21, the device location cannot be identified using DHCP option 82. This is because DHCP option 82 does not apply to the scenario in which MxUs are cascaded. For such a network, LLDP can be used for topology discovery. Specifically, LLDP is used for discovering and creating the link topology information of adjacent devices. The NMS scans the link topology MIBs of all MxUs and their upper-layer devices, and builds the topology of the entire network. The location of an MxU can be identified through the connections between MxUs and between MxUs and their upper-layer devices. In this network scenario, the IP address is also obtained using DHCP. This network scenario applies to OLTs as well.

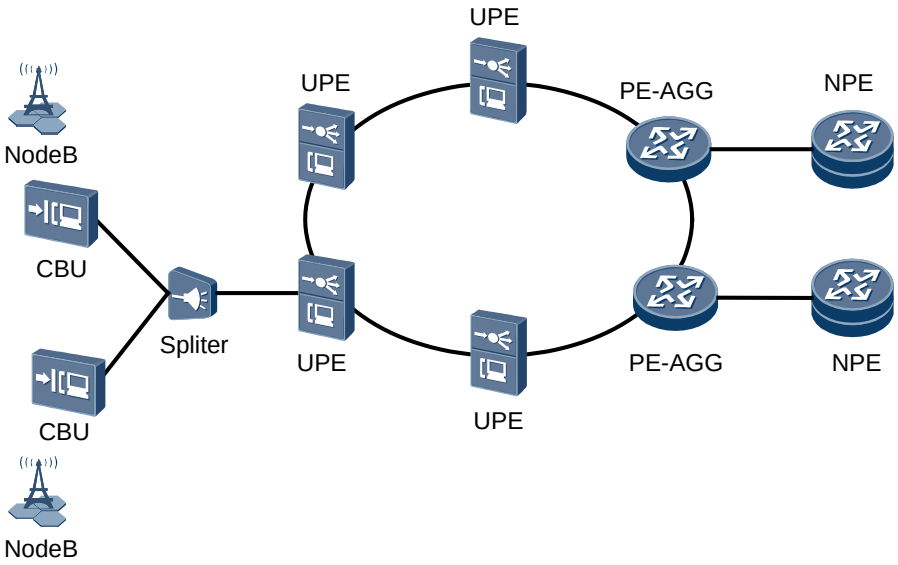
Figure 24-21 Ring network topology



Multipoint-to-Point Connection

As shown in Figure 24-22, the MxU is connected to the upstream device through GPON/EPON. Multiple MxUs are connected to the same PON port, and each MxU is independent of each other in terms of location. In an xPON network, the network topology can be discovered using the automatic discovery function of the ONUs on a PON line, and automatic configuration of ONUs can be implemented using the OMCI protocol. In this scenario, topology discovery between the OLT and MxUs does not require the LLDP protocol. However, LLDP can still be used for topology discovery between OLTs or between an OLT and its upper-layer device.

Figure 24-22 Multipoint-to-point network topology



24.11.5 Configuring LLDP

If you need to manage devices of different vendors and networks, you can configure Link Layer Discovery Protocol (LLDP) on all the devices to obtain the detailed information, such as connections between network topologies and devices. This section describes how to configure LLDP on one device.

Prerequisites

All devices on a network support LLDP so that the network topology can be set up on the U2000.

Procedure

Run the **lldp enable** command to enable the global LLDP function.

Step 1 Run the **lldp port enable** command to enable the port LLDP function.

LLDP must be enabled on all ports that are connected to devices. All LLDP-related configurations take effect only when both global LLDP and port LLDP are enabled.

Step 2 Run the **lldp management address** command to configure the LLDP management address.

Generally, the interface IP address for NMS management is set as the LLDP management address so the management interface IP addresses of devices can be queried using the LLDP topology.

Step 3 (Optional) Configure LLDP parameters.

NOTE
It is recommended to configure the same LLDP parameters on all devices.

- Run the **lldp message-transmission interval** command to configure the interval between sending LLDP notification packets.
- Run the **lldp message-transmission hold-multiplier** command to configure the message hold multiplier. The message hold multiplier is used to calculate the TTL value of a sent packet. The system compares 65535 with hold-value x interval-value + 1, the smaller one is used as the TTL value.
- Run the **lldp restart-delay** command to configure the delay for switching the global or port LLDP status. The delay controls the minimum interval for switching the global or port LLDP status so that the system burden does not increase due to frequently changes of the status.

Step 4 (Optional) Configure LLDP trap function.

- Run the **lldp trap enable** command to enable the LLDP trap sending function.
- Run the **lldp trap interval** command to configure the interval for a port enabled with Link Layer Discovery Protocol (LLDP) to send traps. After the interval is configured, the device will not send a large number of traps to trouble the NMS or subscribers.

Step 5 Query the LLDP information.

- Run the **display lldp local** command to query the local LLDP information, including the system information, LLDP configuration, number of remote neighbors, and ports enabled with LLDP.

- Run the **display lldp neighbor** [**port***frameid/slotid/portid*] command to query the detailed neighbor information about the local device and port relationship between the neighbor device and local device.
- Run the **display lldp neighbor brief** command to query the brief neighbor information.
- Run the **display lldp statistics** command to query the packet statistics of all the ports or a specific port.

Step 6 (Optional) Reset the LLDP information.

- Run the **reset lldp statistics** command to clear the packet statistics of all the ports or a specific port.
- Run the **clear lldp neighbor** command to clear the neighbor information about of all the ports or a specific port.

----End

Example

Assume that:

- Management IP address: 10.10.10.2
- Ports connected to devices: 0/19/0 and 0/19/1
- Other parameters: default values

To configure LLDP and enable LLDP alarming, do as follows:

```
huawei(config)#lldp enable
huawei(config)#lldp enable port 0/19 0,1
huawei(config)#lldp management-address 10.10.10.2
huawei(config)#lldp trap enable
huawei(config)#display lldp local
  System information:
  ChassisId subtype              : macAddress
  ChassisId                     : 00e0-a128-1001
  System name                   : MA5600T
  System description             : Huawei Integrated Access Software
  System capabilities supported    : bridge
  System capabilities enabled      : bridge
  LLDP up time                  : 2013-08-21 20:39:03+08:00

  System configuration:
  LLDP status                   : enabled
  LLDP message TX interval        : 30s
  LLDP message TX hold multiplier   : 4
  LLDP restart delay             : 2s
  LLDP trap interval             : 30s
  LLDP trap enable               : enabled
  Management address             : 10.10.10.2

  Remote neighbors statistics:
  Remote neighbors last change time : 0 day(s),  0h: 0m: 0s
  Remote neighbors added          : 0
  Remote neighbors deleted        : 0
  Remote neighbors dropped        : 0
  Remote neighbors aged           : 0
```

```
Total remote neighbors          : 0

Port information:

Port                          : 0 /19/0
PortId subtype                  : interfaceName
PortId                        : ethernet0/19/0
Port description                : Huawei-MA5600-V800R013-ETHERNET
LLDP enable status              : enabled TX-RX
Total remote neighbors          : 0

Port                          : 0 /19/1
PortId subtype                  : interfaceName
PortId                        : ethernet0/19/1
Port description                : Huawei-MA5600-V800R013-ETHERNET

LLDP enable status              : enabled TX-RX
Total remote neighbors          : 0

Total port number               : 2
```

24.11.6 Reference Standards and Protocols

The following lists the reference standards and protocols of the LLDP feature:

- IEEE 802.1ab-2009
- IEEE 802.1ab-2005 (serves as the reference for the destination MAC address part of the LLDP feature)

24.12 Alarm and Event Management

This topic covers the overview and working principle of alarm and event management.

24.12.1 Introduction

Definition

Alarm and event management mainly involves recording and setting alarms and events and collecting their statistics.

Purpose

Alarm and event management facilitates carriers in performing routine maintenance on the device, locating device faults, and restoring the services provided for users quickly after the services become abnormal.

24.12.2 Principle

The alarm and event management refers to recording and setting the alarms and events and collecting statistics of the alarms and events. The maintenance engineers maintain the device through the alarm and event management so that the device works effectively.

After an alarm or event is generated, the system broadcasts the alarm or event to the terminals, mainly including the Network Management System (NMS) and Command Line Interface (CLI) terminals. Currently, the system supports storing 1000 history alarms and 800 history events.

The severity level of an alarm or event can be critical, major, minor, or warning. Although an alarm or event has a default severity level, this severity level can be adjusted according to actual conditions. The contents of an alarm or event include name, parameters (including subrack, slot, and port information), description, possible causes, and handling suggestions.

When an alarm is generated, the system implements the jitter-proof function of the alarm to prevent the misreporting of the alarm. To be specific, the alarm is reported only after a specified period expires after the alarm status changes (the specified period ranges from 1s to 60s and is 10s by default). If the alarm status recovers within the specified period, the alarm is not reported.

The alarm statistics function is used to collect the statistics of alarms within a specified period. This helps to locate system faults.

Alarm correlation refers to associating related alarms. When alarms are in the parent-child relations, the system automatically filters related child alarms if the parent alarm is generated.

With the alarm and event filtering function, the user can configure the filtering conditions so that the system reports only the alarms and events that pass the filtering. In this way, the user can concentrate on the important and specified alarms and events. The alarms and events can be filtered according to their ID, severity level, and type.

24.13 Relevant Standards and Protocols

SNMP

The following lists the reference standards and protocols of device security:

1. SNMPv1
   - RFC1157: Simple Network Management Protocol (SNMP)
2. SNMPv2c
   - RFC1905: Protocol Operations for Version 2 of the Simple Network Management Protocol (SNMPv2)
3. SNMPv3
   - RFC2570: Introduction to Version 3 of the Internet-standard Network Management Framework
   - RFC2571: An Architecture for Describing SNMP Management Frameworks
   - RFC2572: Message Processing and Dispatching for the Simple Network Management Protocol (SNMP)
   - RFC2573: SNMP Applications
   - RFC2574: User-based Security Model (USM) for version 3 of the Simple Network Management Protocol (SNMPv3)
   - RFC2575: View-based Access Control Model (VACM) for the Simple Network Management Protocol (SNMP)

SSH

Encryption for remote management connection:

- RFC4254: The Secure Shell (SSH) Connection Protocol
- RFC4253: The Secure Shell (SSH) Transport Layer Protocol
- RFC4252: The Secure Shell (SSH) Authentication Protocol
- RFC4251: The Secure Shell (SSH) Protocol Architecture

Encryption for file transfer:

None

User Name/Password Management

None

ANCP

- IETF RFC6320/draft00/draft01/draft02: Protocol for Access Node Control Mechanism in Broadband Networks

Remote Connection Security

None

Log Management

None

Version and Data Management

None

Alarm and Event Management

None

25 Service Overload Control

About This Chapter

This topic provides the definition, purpose, and principle of service overload control.

25.1 Introduction

Definition

Overload control (OLC) is a mechanism that prevents exhaustion of system resource such as CPU resources. It protects equipment from service interruption or NMS unreachability triggered by overload of CPU or other resources in the event of heavy traffic. OLC also ensures to a certain extent the quality of high priority services (such as emergency calls) when the system is overloaded.

Purpose

On the live network, the CPU usage or service resources on the access equipment may be overloaded in any of the following conditions:

- Protocol packet flooding
- Alarm packet flooding
- Burst traffic due to a large number of concurrent online users
- Frequent data loading, query, or save operations in the system

The MA5600T/MA5603T/MA5608T provides the OLC feature to ensure that the system is able to successfully process services in any of the above-mentioned conditions.

Benefits

Benefits to carriers

The OLC-enabled MA5600T/MA5603T/MA5608T is able to filter and control the packets sent to the CPU to defend the system against malicious attacks and instantaneous service overload, improving device security and reliability.

25.2 Principle

The packets sent to the CPU must be specified with priorities. The packets include internal management packets, network topology management packets, and service (voice and broadband services) packets.

The system may have the following packets:

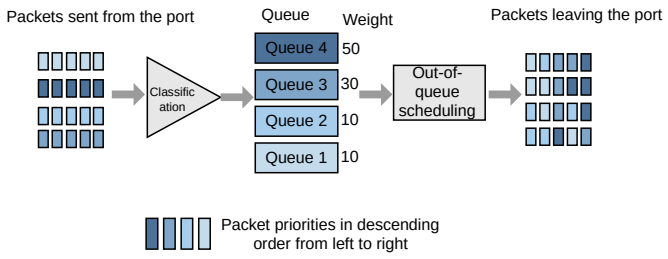
- Internal management packets, including inter-board handshake packets, upper-layer protocol packets, and packets of loading tasks
- Link-layer network management packets such as MSTP and LACP packets
- Protocol packets such as routing protocol, BFD, and ETH OAM packets
- SNMP, ANCP, and NTP packets
- VoIP, IPTV, and private line service packets

To differentiate packets with different priorities, the OLC feature must support different priorities for different queues. It employs the weighted round robin (WRR), strict priority (SP), and token bucket algorithms for queue scheduling.

WRR

Figure 25-1 illustrates the principle of the WRR algorithm.

Figure 25-1 WRR algorithm

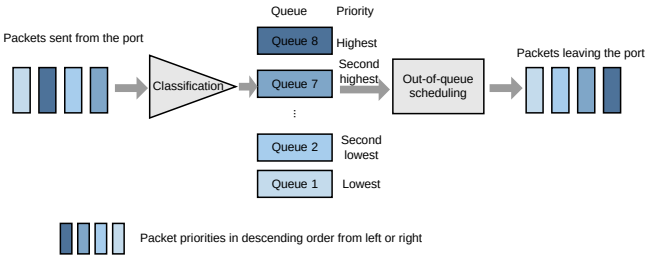


In the WRR algorithm, a weight value is assigned to each queue based on round robin (RR) and a counter is maintained for each queue. During each RR, queues whose counter value is not zero are allowed to send a packet. The initial value of the counter is the weight of a queue. Each time a packet is sent, one is subtracted from the counter value regardless of whether the packet is successfully scheduled or not. When the counter values of all the queues become zero, the counter values are reset to their initial values. The WRR algorithm achieves fairness among queues and smoothly schedules outbound services.

SP

Management packets, voice packets, and important broadband protocol packets are scheduled using the SP algorithm. Figure 25-2 illustrates the principle of the SP algorithm.

Figure 25-2 SP algorithm

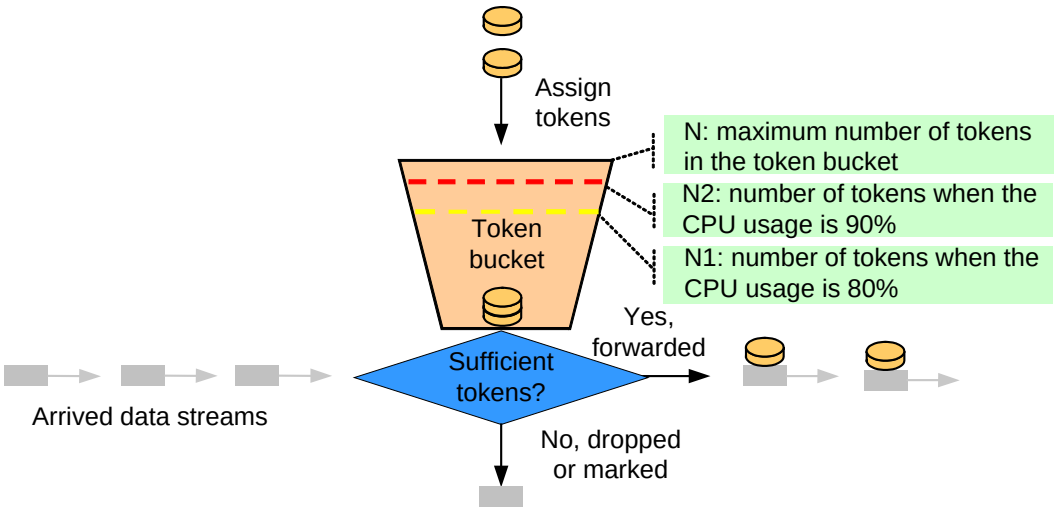


In SP queue scheduling, packets are sent in descending order of queue priorities. When a queue with the highest priority is empty, the packets in the queue with the second highest priority are sent. In this way, packets for critical services in a queue with a higher priority are sent first, and packets of non-critical services (such as email service) in a queue with a lower priority are sent only when the network is idle and the critical services have been processed.

Token Bucket

The MA5600T/MA5603T/MA5608T provides a token bucket. The system assigns a certain number (at a leakage rate of a) of tokens to the token bucket every second. The number is determined by the system processing capability assuming that the CPU usage is 80%. If a packet arrives when the tokens in the token bucket are used up, the packet will be dropped or marked. Figure 25-3 shows the principle of the token bucket algorithm.

Figure 25-3 Token bucket algorithm



The token bucket algorithm involves the following key parameters:

- Leak rate (a): number of tokens in the token bucket. The leak rate is dynamically adjusted according to the CPU usage. When the CPU usage is higher than the preset CPU threshold, the leak rate is lowered to slow down the arrival rate of packets. When the CPU usage is lower than the preset CPU threshold, the leak rate is raised to speed up packet processing.

- Token bucket threshold (N): default system capacity (default: 1000; unit: packet, or token number.) This threshold indicates the standard packet processing capability of the system assuming that the CPU usage is 100%. The value range cannot be modified.
- Target CPU usage threshold (T1): target (level-1) CPU usage threshold, the upper threshold of the CPU usage (range: 70%-99%; default value: 80%). The corresponding leak number N1 = N x T1. When the CPU usage exceeds T1 (80%), the system starts lowering the leak rate.
- Level-2 CPU usage threshold (T2): range: 71%-100%; default: 90%. The corresponding water level N2 = N x T2. When the CPU usage exceeds 90%, the system lowers the leak rate at a faster pace. Therefore, packets are discarded at a fast pace.
- Overload control startup threshold (T3): is same as the target CPU usage threshold (T1). The overload control is enabled when the CPU usage reaches the overload control startup threshold. If overload control is disabled, no packet is randomly dropped.
- Overload control restoration threshold (T4): is lower than the overload control startup threshold. The value of T3 is calculated based on the following formula: T4 = T3 - 20%

NOTE

If the overload control restoration threshold is the same as the overload control startup threshold, the overload control is enabled and disabled repeatedly when the CPU usage changes around the overload control startup threshold. This causes a flapping, which greatly affects CPU usage stability.

- Adjustment factor (S): adjustment step (range: 1-1000; default: 10). The smaller the adjustment factor, the quicker the upshift or downshift of the leak rate, and the larger the leak rate jitter. Reversely, the slower the upshift or downshift of the leak rate, the smaller the leak rate jitter.

Relationship of the three Algorithms

When implementing WRR scheduling, the system also determines whether to read packets from the current queue according to the overload status of the token bucket. The system employs different algorithms to schedule packets sent to the CPU: employs SP for queues of management packets and VoIP packets, employs token bucket for queues of voice packets (the leak rate is dynamically assigned based on the CPU usage), and employs WRR for other queues. In summary, the system uses different algorithms (token bucket, WRR, and SP) to schedule packets in case of system overload so that:

1. Different services in the same queue can be fairly processed to ensure that the burst of a certain type of service packets does not affect other types of service packets.
2. The non-protocol packets are controlled by the system-level OLC feature to ensure task fairness and provide basic guarantee for services such as upgrade and user dialup.

26 System Security

About This Chapter

This topic provides the plan security schemes based on site requirements, and describes the sub-features of this feature.

26.1 Security Scheme Planning

Plan security schemes based on site requirements to protect the system against attacks from malicious users.

Table 26-1 System security schemes

Vulnerability	Security Scheme	Suggestion
Malicious users send a large number of protocol packets to attack the MA5600T/MA5603T/MA5608T. In this case, the MA5600T/MA5603T/MA5608T fails to process service requests from common users.	26.2 DoS Anti-Attack	Use this security scheme during deployment.
Malicious users send Internet Control Message Protocol (ICMP) or IP packets to the MA5600T/MA5603T/MA5608T. As a result, MA5600T/MA5603T/MA5608T resources are exhausted and the MA5600T/MA5603T/MA5608T may malfunction. For example: • Malicious users send a large number of ping packets to request responses from the	• 26.3 IP or ICMP Anti-Attack on the User Side • 26.3 IP or ICMP Anti-Attack on the User Side	Use this security scheme when the following conditions are met: • The MA5600T/MA5603T/MA5608T works at Layer 3. • The destination address of the ICMP or IP

Vulnerability	Security Scheme	Suggestion
MA5600T/MA5603T/MA5608T. As a result, the MA5600T/MA5603T/MA5608T becomes overloaded. • Malicious users may find system vulnerabilities by pinging or telnetting the MA5600T/MA5603T/MA5608T and then initiate attacks. **NOTE** It is recommended to log in to the MA5600T/MA5603T/MA5608T through Secure Shell (SSH).		packets is not planned to be the system IP address of the MA5600T/MA5603T/MA5608T.
Attackers forge the IP addresses of authorized users to attack networks by setting source route options. In this case, the MA5600T/MA5603T/MA5608T fails to process service requests from common users.	26.4 Source Route Filtering	Use this security scheme during deployment.
Unauthorized terminal or management users communicate with the MA5600T/MA5603T/MA5608T.	26.5 Firewall	Use this security scheme during deployment.

The following common improper configurations affect system security:

- Devices are managed using a public IP address. Access rights are not limited when an access control list (ACL) is configured. As a result, the network is prone to attacks.

  Preventive methods or measures:

  - Use a private IP address to manage devices.
  - When configuring the ACL, grant minimum rights to users within the minimum range.
  - Configure a permitted IP address segment, and add only desired management IP addresses to the IP address segment. Users with an IP address out of the IP address segment cannot log in to the management port on the MA5600T/MA5603T/MA5608T.
- Packets received by the management port on the MA5600T/MA5603T/MA5608T are not controlled. As a result, if a large number of packets are sent to the MA5600T/MA5603T/MA5608T, the MA5600T/MA5603T/MA5608T becomes too busy to provide services.

  Preventive methods or measures: Run the **firewall packet-filter** command to apply a firewall filtering rule to the management port.

Table 26-2 lists the default settings of system security.

Table 26-2 Default settings of system security

Parameter	Default Setting
Firewall blacklist	Disabled
DoS anti-attack	Disabled
ICMP anti-attack	Disabled
IP address anti-attack	Disabled
Source route filtering	Disabled

26.2 DoS Anti-Attack

The DoS anti-attack feature enables the system to receive or drop protocol packets sent by users based on specified limitations. When hit by a DoS attack, the system is incapable of responding to user service packets. This feature prevents attacks on the system initiated by malicious users who send a large number of protocol packets.

26.2.1 What Is DoS Anti-Attack

Definition

A denial of service (DoS) attack is initiated by malicious users using a large number of protocol packets. When hit by a DoS attack, the system becomes incapable of responding to user service packets, which prevents the system from serving common users.

The DoS anti-attack feature limits the rate of incoming protocol packets and performs blacklist management for malicious users that initiate DoS attacks. It achieves the following:

- Protects carrier networks against attacks so that their access devices function properly and securely.
- Enhances service security for terminal users so that users can enjoy stable and secure services.

26.2.2 Principles

Deny of service (DoS) anti-attack on the MA5600T/MA5603T/MA5608T safeguards CPU resources using blacklists and packet processing policies for DoS anti-attacks.

- Blacklist for DoS anti-attacks: The administrator of the MA5600T/MA5603T/MA5608T adds a port or GPON encapsulation mode (GEM) port where a DoS attack is initiated to the blacklist and forces the users in the blacklist to go offline.
- Packet processing policies for DoS anti-attacks: If the MA5600T/MA5603T/MA5608T receives a protocol packet at a rate higher than the rate threshold, the MA5600T/MA5603T/MA5608T limits the packet rate or discards the packet according to the configured packet processing policy.

NOTE

1. The MA5600T/MA5603T/MA5608T limits the rate of protocol packets but does not report them to the blacklist if the firewall blacklist function is disabled and the packet processing policy for DoS anti-attacks is **deny** (protocol packets are discarded).
2. Packet processing policies for DoS anti-attacks take effect only after the DoS anti-attack blacklist function is enabled.
3. If a blacklist is generated, the blacklist is deleted after the DoS anti-attack policy is switched, for example, from **deny** to **permit**. Then, the system performs a DoS anti-attack detection again.

Blacklist for DoS Anti-Attacks

The MA5600T/MA5603T/MA5608T maintains a DoS attack blacklist. After DoS anti-attack blacklist is enabled, maintenance personnel can detect a DoS attack promptly through a DoS attack alarm, such as **0x29000008 A DoS attack occurs on the user port**, or by running the **display security dos-blacklist** command. Them, the maintenance personnel isolate or even disconnect the malicious user. For example, the maintenance personnel can deactivate the port to force the malicious user to go offline.

1. **Add a port to the blacklist.**

   If the number of DoS attacks that are consecutively detected by the MA5600T/MA5603T/MA5608T on a port is the same as that of detection times, the MA5600T/MA5603T/MA5608T adds the port to the blacklist and reports a DoS attack alarm. In this case, DoS attack occurred on the port.

2. **Update the blacklist.**

   The MA5600T/MA5603T/MA5608T continuously checks the DoS attack activities of the members in the blacklist.

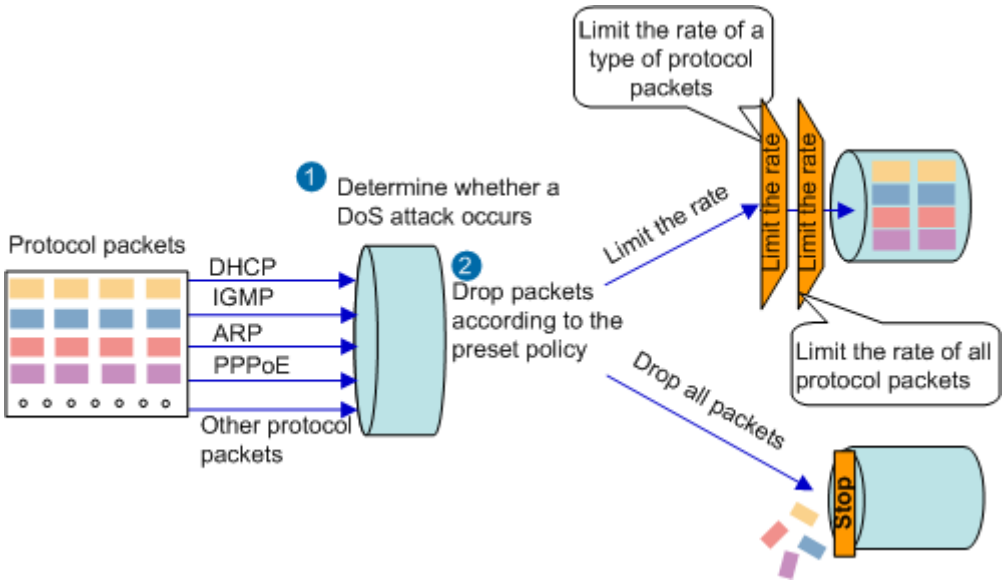
   - GPON boards: Upon adding a GEM port to the blacklist, the MA5600T/MA5603T/MA5608T detects a DoS attack again.
   - xDSL and P2P interface boards: The MA5600T/MA5603T/MA5608T detects a DoS attack again after adding a port to the blacklist 3 minutes later.

   If no DoS attack occurs on the port during the consecutive detection for DoS attacks, the MA5600T/MA5603T/MA5608T deletes the port from the blacklist

Packet Processing Policies for DoS Anti-Attacks

Figure 26-1 shows the flow of limiting the rate of protocol packets when a DoS attack occurs.

Figure 26-1 Flow of limiting the rate of protocol packets when a DoS attack occurs



1. The MA5600T/MA5603T/MA5608T supports four rate thresholds for IGMP, DHCP, ARP, and PPPoE packets, and one total rate threshold for all protocol packets. In a detection period, if the rate of protocol packets of a type exceeds a threshold, the MA5600T/MA5603T/MA5608T determines that the port encounters a DoS attack.
2. The MA5600T/MA5603T/MA5608T supports the following two packet processing policies for the user initiating a DoS attack:
   - Permits protocol packet sending and limit the packet rate (packet processing policy **permit**): When a DoS attack occurs, the MA5600T/MA5603T/MA5608T limits the rate of only received protocol packets so that this rate is lower than the preset threshold. The rate of IGMP, DHCP, ARP, and PPPoE packets is limited and the total rate of all packets is limited.
   - Discards protocol packets (packet processing policy **deny**): When a DoS attack occurs, the MA5600T/MA5603T/MA5608T discards all the received protocol packets.

NOTE
The MA5600T/MA5603T/MA5608T discards all protocol packets received on the port if either rate of IGMP, DHCP, ARP, or PPPoE packets exceeds the preset threshold.

If a blacklist is generated after a DoS attack and the packet processing policy is set to **deny**:

- For ARP and ND packets, the MA5600T/MA5603T/MA5608T forwards the packets to the uplink port and discards the packets destined for the CPU.
- For other protocol packets, the MA5600T/MA5603T/MA5608T discards the packets destined for the uplink port and CPU.

An example is provided here to describe the packet processing. After Policy Information Transfer Protocol (PITP) or MAC address anti-spoofing is enabled, the MA5600T/MA5603T/MA5608T sends PPPoE discovery packets to the CPU. If a DoS attack occurs and the packet processing policy is set to **deny**, the MA5600T/MA5603T/MA5608T discards the PPPoE discovery packets sent from the port that is contained in the blacklist.

26.2.3 Configuring DoS Anti-attack

The configuration of denial of service (DoS) anti-attack and processing policies for protocol packets prevents malicious users from sending a great number protocol packets to attack the MA5600T/MA5603T/MA5608T.

Context

Table 26-3 lists system security schemes.

Table 26-3 System security schemes

Vulnerability	Security Scheme	Suggestion
Malicious users send a large number of protocol packets to attack the MA5600T/MA5603T/MA5608T. In this case, the MA5600T/MA5603T/MA5608T fails to process service requests from common users.	26.2 DoS Anti-Attack	Use this security scheme during deployment.

Procedure

Run the **security anti-dos enable** command to enable DoS anti-attack.

Step 1 Run the **security anti-dos control-packet policy** command to configure a processing policy for protocol packets when a DoS attack occurs.

NOTICE

The processing policy for protocol packets takes effect only after DoS anti-attack is globally enabled.

Step 2 Run the **security anti-dos control-packet rate** command to configure the rate threshold for sending protocol packets to the CPU.

----End

Example

The following configurations are used as an example to globally enable DoS anti-attack:

- Processing policy for protocol packets when a DoS attack occurs: adding the information of the attacking port to the DoS blacklist
- Rate threshold on port 0/2/1 for sending protocol packets to the CPU: 20 pps

```
huawei(config)#security anti-dos enable
huawei(config)#security anti-dos control-packet policy permit
```

```
huawei(config)#security anti-dos control-packet rate 0/2/1 20
```

26.3 IP or ICMP Anti-Attack on the User Side

The feature of IP or ICMP anti-attack on the user side enables the MA5600T/MA5603T/MA5608T to identify and discard any IP or Internet Control Message Protocol (ICMP) packet sent from end users whose destination IP address is the same as the system IP address. Therefore, this feature allows the MA5600T/MA5603T/MA5608T to avoid IP or ICMP attacks initiated from the user side. The system IP address of the MA5600T/MA5603T/MA5608T includes the management IP address and the IP address of the Layer 3 interface.

26.3.1 What Are IP/ICMP Attacks from the User Side

IP Attacks from the User Side

The destination IP addresses of packets sent from common users are typically different from the system IP address of an access device (excluding the special planning of some carriers). Malicious users forge IP packets with the destination IP address set to the system IP address and attack the access device with the IP packets. During a common IP attack, malicious users send a large number of packets to request responses from the access device. As a result, the access device becomes overloaded and fails to process service requests from common users. IP attacks can be considered as a type of denial of service (DoS) attack.

To avoid IP attacks from malicious users, the access device can identify and discard any received IP packets whose destination IP address is the same as the system IP address.

ICMP Attacks from the User Side

The Internet Control Message Protocol (ICMP) is a sub-protocol of the TCP/IP protocol suite. It is used for transmitting control messages, such as ping and route tracing messages, between IP hosts and routers. During fault locating, ICMP packets can be sent from the peer device to the access device to check network connectivity and route reachability. ICMP attacks can be avoided in two ways.

Only the upper-layer device or cascading device is allowed to ping the access device. The user terminal is therefore not allowed to ping the access device, preventing the user terminal from initiating attacks after successfully pinging and detecting the access device.

The access device can identify and discard any received ICMP packet whose destination IP address is the same as the system IP address to avoid ICMP attacks from malicious users.

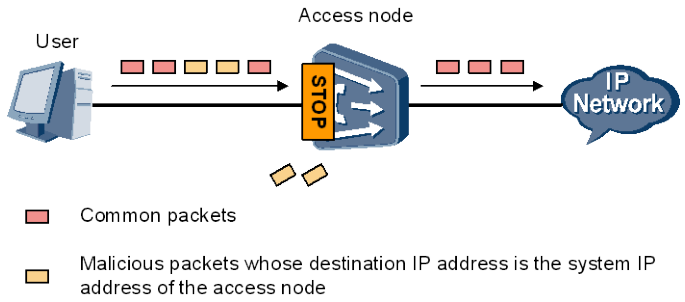
26.3.2 Principles of User-side IP/ICMP Anti-Attacks

The MA5600T/MA5603T/MA5608T can identify and discard any IP or Internet Control Message Protocol (ICMP) packets whose destination IP address is the same as the system IP address to avoid IP or ICMP attacks from users, as shown in Figure 26-2.

NOTE

The system IP address of the MA5600T/MA5603T/MA5608T includes the management IP address and the IP address of the Layer 3 interface.

Figure 26-2 IP or ICMP anti-attack



In the preceding figure, the MA5600T/MA5603T/MA5608T:

- Forwards common packets.
- Discards any malicious packet whose destination IP address is the same as the system IP address.

26.3.3 Configuring ICMP or IP Address Anti-attack

The configuration of Internet Control Message Protocol (ICMP) or IP address anti-attack prevents malicious users from sending ICMP or IP packets to exhaust system resources and ensures the normal running of access devices.

Context

Table 26-4 lists system security schemes.

Table 26-4 System security schemes

Vulnerability	Security Scheme	Suggestion
Malicious users send Internet Control Message Protocol (ICMP) or IP packets to the MA5600T/MA5603T/MA5608T. As a result, MA5600T/MA5603T/MA5608T resources are exhausted and the MA5600T/MA5603T/MA5608T may malfunction. For example: • Malicious users send a large number of ping packets to request responses from the MA5600T/MA5603T/MA5608T. As a result, the MA5600T/MA5603T/MA5608T becomes overloaded. • Malicious users may find system vulnerabilities by pinging or telnetting the	• 26.3 IP or ICMP Anti-Attack on the User Side • 26.3 IP or ICMP Anti-Attack on the User Side	Use this security scheme when the following conditions are met: • The MA5600T/MA5603T/MA5608T works at Layer 3. • The destination address of the ICMP or IP packets is not planned to be the system IP address of the MA5600T/MA5603T/MA5608T.

Vulnerability	Security Scheme	Suggestion
MA5600T/MA5603T/MA5608T and then initiate attacks. **NOTE** It is recommended to log in to the MA5600T/MA5603T/MA5608T through Secure Shell (SSH).		

Procedure

- Run the **security anti-icmpattack enable** command to enable ICMP anti-attack.
- Run the **security anti-ipattack enable** command to enable IP anti-attack.

----End

Example

The following configurations are used as an example to enable ICMP and IP address anti-attack so that users cannot send ICMP and IP packets to the access device:

```
huawei(config)#security anti-icmpattack enable
huawei(config)#security anti-ipattack enable
```

26.4 Source Route Filtering

Source route filtering enables the system to identify and drop IP packets with source route options, and therefore to prevent malicious users from attacking networks using source route options.

26.4.1 Why Source Route Filtering Is Required

Source route filtering is intended to resolve issues caused by source route options.

Source Route Option

The packet transmission route of an IP packet can be defined in the IP packet header on an IPv4 network by a strict source route option or a loose source route option.

- If a packet carries a strict source route option, the packet must be forwarded exactly (hop by hop) according to the routers specified by the option.
- If a packet carries a loose source route option, the packet is forwarded according the routers specified by the option, but the packet may traverse other routers between two specified routers.

Functions of Source Route Option

Users can specify the route or part of the route for packets by adding source route options into the packets. Packets can therefore be selectively sent to different addresses. Source route

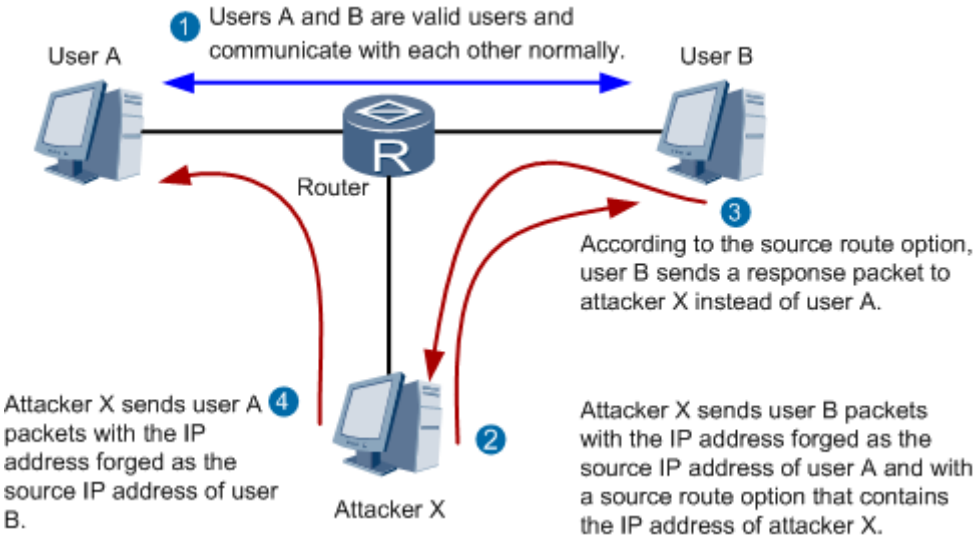
options can be used to test the throughput rate of a network or to transmit data over a trustworthy network. For example, for an IP packet to traverse the routers R1, R2, and R3, users can specify the interface addresses of the three routers in the strict source route option. Then, the IP packet traverses the three routers regardless of the route tables on the routers.

When the peer device receives a packet with a source route option, it responds with a packet with a source route option. The router order in the source route option of the sent packet is the opposite of that in the source route option of the received packet. The response packet can therefore be forwarded along the same route.

Issues Brought by Source Route Option

In a secure network, users manage the directions of data stream flows by specifying packet forwarding routes in the source route option. However, in an insecure network, malicious users may use source route options to attack the network and intercept authorized users' communication data. Figure 26-3 shows how malicious users attack networks using source route options.

Figure 26-3 How malicious users attack networks using source route options



1. Users A and B are authorized users and communicate with each other normally.
2. Attacker X sends a packet with a source route option to user B. The source IP address of the packet sent by attacker X is forged as the IP address of user A, and the source route option contains the IP address of attacker X.
3. User B receives the packet sent from attacker X and believes that the packet is sent from user A because the packet carries the IP address of user A as the source IP address. According to the source route option in the received packet, user B intends to send a response packet to user A, but actually sends the response packet to attacker X.
4. Attacker X receives the packet sent from user B, forges the IP address of user B, and sends a packet to user A. In addition, attacker X can also modify the source route option so that the packet is forwarded to user A through the specified route. Attacker X can therefore conceal their actual location.

Solution

When malicious users attack networks, they use source route options as an auxiliary method of IP address spoofing. The following describes how to protect access devices from malicious users' attacks.

- Source route filtering: Filter out IP packets that are sent by the user and carry source route options.
- IP anti-spoofing: Prevent malicious users from forging the IP addresses of authorized users.

26.4.2 Configuring Source Route Filtering

The configuration of source route filtering prevents malicious users from forging authorized IP addresses to attack the network and ensures the network can process service requests from common users.

Context

Table 26-5 lists system security schemes.

Table 26-5 System security schemes

Vulnerability	Security Scheme	Suggestion
Attackers forge the IP addresses of authorized users to attack networks by setting source route options. In this case, the MA5600T/MA5603T/MA5608T fails to process service requests from common users.	26.4 Source Route Filtering	Use this security scheme during deployment.

Procedure

- Run the **security source-route enable** command to enable source route filtering.

  The source route filtering function filters out the packets that carry routing information and are sent to Layer 3.

----End

Example

The following configurations are used as an example to enable source route filtering:

```
huawei(config)#security source-route enable
```

26.5 Firewall

A firewall is an advanced access-control mechanism deployed between network security zones to control access to the network by implementing security policies.

26.5.1 Why Firewall Is Required

Common Internet security threats can be classified as follows:

- Unauthorized use: Resources are used without authorization. For example, attackers gain access to a computer system and use resources by guessing a user account and password combination.
- Denial of service (DoS): Attackers exploit vulnerabilities of network protocol implementation to initiate attacks or maliciously exhaust resources of the attacked object. A DoS attack is an attempt to stop the target object from providing services or resources. For example, attackers send a large amount of data packets or deformed packets to a server to request for connections or replies, overloading the server so much that the server cannot process service requests from common users.
- Data tampering: Attackers modify, delete, delay, or realign system data or message flows, or insert fake messages to compromise data consistency.
- Information theft: Attackers do not invade the target system, but sniff it to steal important data or information.

A firewall monitors and determines whether data flows are allowed to enter an access device by analyzing data packets. It protects internal networks against unauthorized or unauthenticated access and attacks from external networks.

The MA5600T/MA5603T/MA5608T filters data packets using the four firewall techniques listed in the following table.

Table 26-6 Firewall techniques supported by the MA5600T/MA5603T/MA5608T

Technique	Function	Feature
Firewall blacklist	A firewall blacklist filters data packets by source IP address.	Matching source IP addresses against a blacklist is simple, and packets can be quickly filtered. However, because data packets are filtered by only one rule, this process lacks flexibility.
Firewall blacklist and advanced access control list (ACL) rules	The combination of a firewall blacklist and advanced ACL rules enables the MA5600T/MA5603T/MA5608T to further filter packets by advanced ACL rules.	Data packets are filtered based on a firewall blacklist and advanced ACL rules. The filter rules can be flexibly configured.
ACL-based packet filtering firewall	An ACL-based packet filtering firewall verifies data packets at the network layer and forwards or denies	Advantage: This technique supports more flexible configurations and better filtering capabilities than

Technique	Function	Feature
	them according to the security policy.	firewall blacklist. Disadvantages: • The packet filtering performance deteriorates sharply as the ACL complexity increases. • The MA5600T/MA5603T/MA5608T does not check the session status or analyze any data, and is vulnerable to IP spoofing attacks.
Unauthorized login prevention	The MA5600T/MA5603T/MA5608T prevents unauthorized logins by setting the IP address segments permitted by denied by the firewall for specified protocol types.	N/A

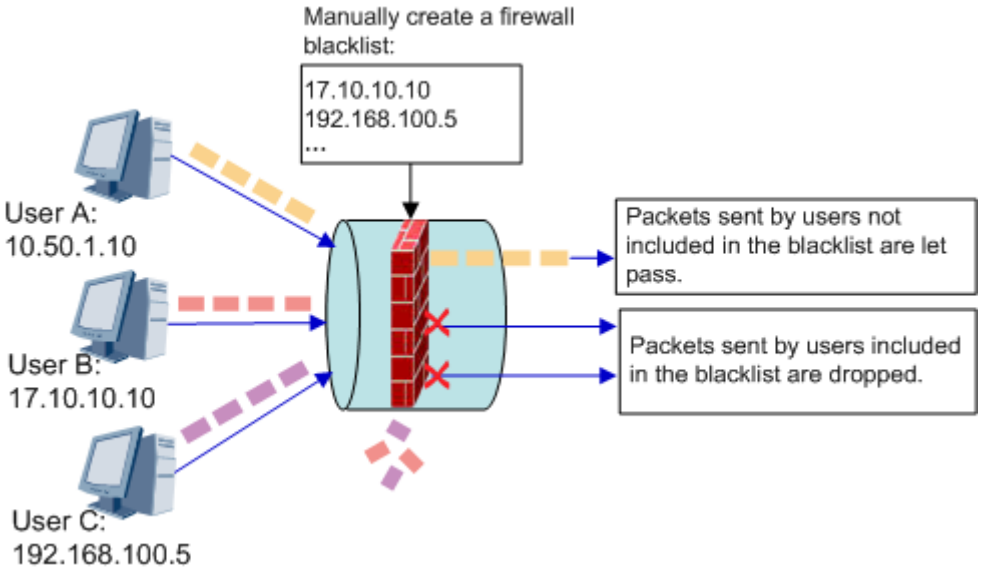
26.5.2 Firewall Filtering

Firewall Blacklist

A firewall blacklist quickly filters packets by source IP address, dropping any unwanted packets that originate from specified IP addresses.

Figure 26-4 shows how the MA5600T/MA5603T/MA5608T implements the firewall blacklist feature.

Figure 26-4 Implementation of firewall blacklist



Users configure a firewall blacklist by running commands. The system then performs the following operations before the firewall blacklist expires:

- Permits access for packets whose IP addresses are not included in the blacklist.
- Drops packets whose IP addresses are included in the blacklist.

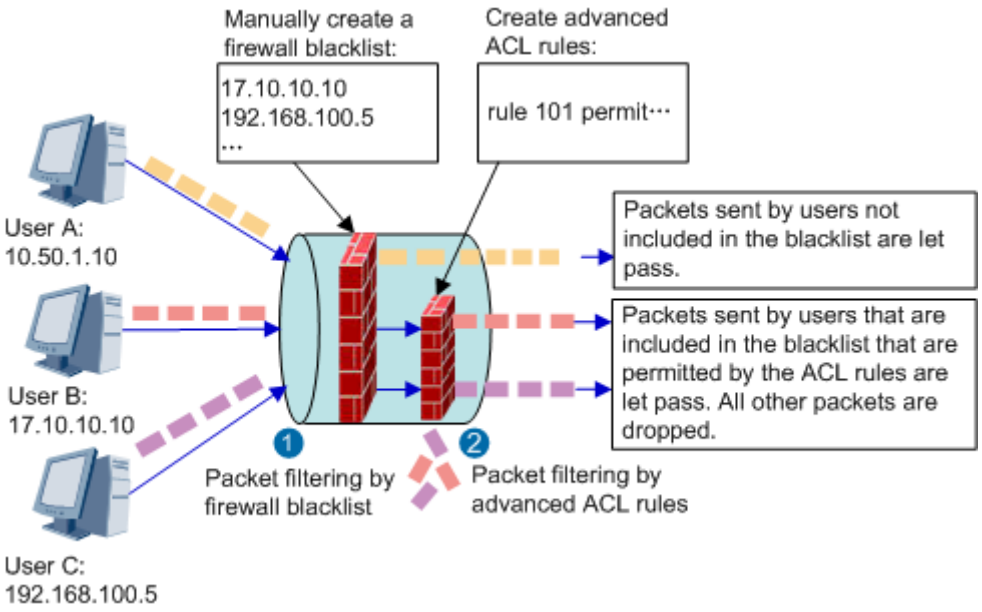
NOTE

- Users can specify the valid duration (aging time) of an IP address in the firewall blacklist. When the duration expires, the IP address is removed from the blacklist. If users do not specify the aging time, the IP address is permanently included in the blacklist unless manually deleted.
- A blacklist entry added to a blacklist takes effect only after the firewall blacklist feature is enabled.

Firewall Blacklist Combined with Advanced ACL Rules

The combination of a firewall blacklist and advanced ACL rules enables the system to filter data packets at a finer grain. Figure 26-5 shows how the system filters data packets based on this combination.

Figure 26-5 Implementation of a firewall blacklist combined with advanced ACL rules



The system filters packets at two levels:

1. Filters packets by blacklist. Users configure a firewall blacklist by running commands. The system then performs the following operations before the firewall blacklist expires:
   - Permits access for packets whose IP addresses are not included in the blacklist.
   - Filters packets whose IP addresses are included in the blacklist according to the advanced ACL rules.

NOTE

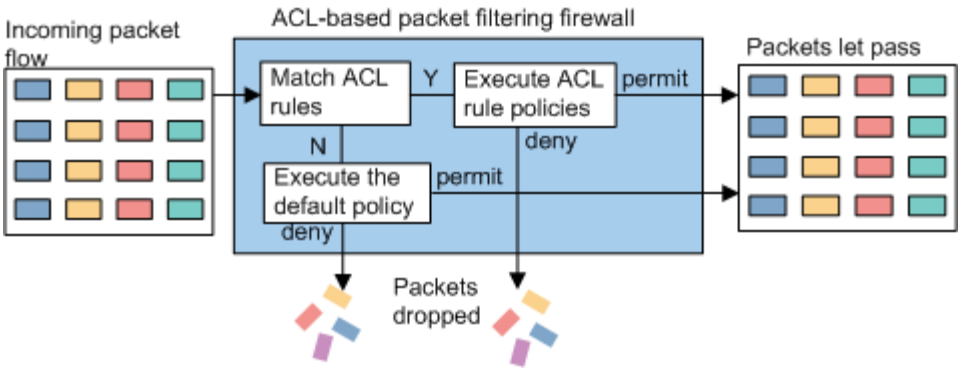
- Users can specify the valid duration (aging time) of an IP address in the firewall blacklist. When the duration expires, the IP address is removed from the blacklist. If users do not specify the aging time, the IP address is permanently included in the blacklist unless manually deleted.
- A blacklist entry added to a blacklist takes effect only after the firewall blacklist feature is enabled.

2. Filters packets by advanced ACL rules. To use advanced ACL rules to filter packets whose IP addresses are included in the firewall blacklist, the system performs the following:
   - Drops packets that match ACL deny rules.
   - Permits packets that match ACL permit rules, regardless of whether the source IP addresses of these packets are included in the firewall blacklist.
   - Drops packets that carry the IP addresses included in the firewall blacklist but do not match any ACL rule.

ACL-based Firewall Filtering

Firewall filtering based on an access control list (ACL) checks data packets at the network layer and forwards or denies them according to the security policy.

Figure 26-6 Implementation of ACL-based firewall filtering



The MA5600T/MA5603T/MA5608T enabled with ACL-based firewall filtering filters data packets according to pre-configured basic or advanced ACL rules.

- A basic ACL rule is configured based on a Layer 3 source IP address. The MA5600T/MA5603T/MA5608T analyzes and processes data packets according to the ACL rule.
- When an advanced ACL rule applies, the MA5600T/MA5603T/MA5608T classifies traffic according to the following factors:
  - Protocol type
  - Source IP address
  - Destination IP address
  - Source port number (source port of UDP or TCP packets)
  - Destination port number (destination port of UDP or TCP packets)
  - Type precedence value of ICMP packets (precedence field of a data packet)
  - Type of service (ToS) value (ToS field of a data packet)
  - Differentiated services code point (DSCP) value (DSCP field of a data packet)

For details about and implementation of ACLs, see 14.10 ACL.

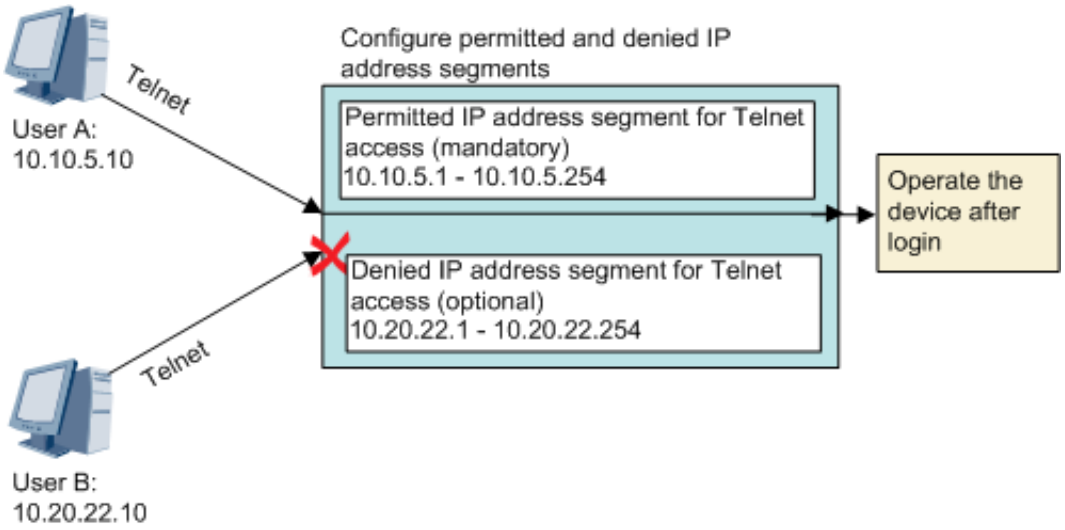
Permitted/Denied IP Address Segment

The MA5600T/MA5603T/MA5608T supports the configuration of an IP address segment permitted or denied by the firewall for specified protocol types to prevent logins from unauthorized IP address segments. This helps secure the system.

When a user attempts to log in to the MA5600T/MA5603T/MA5608T through Telnet, Secure Shell (SSH), or Simple Network Management Protocol (SNMP), the MA5600T/MA5603T/MA5608T checks whether the IP address of the user is included in the permitted or denied IP address segment and permits or denies the login attempt accordingly. Figure 26-7 shows the process of permitting or denying Telnet access.

NOTE
It is recommended to log in to the MA5600T/MA5603T/MA5608T through SSH.

Figure 26-7 Process of permitting or denying Telnet access



- The permitted IP address segment is a mandatory configuration item. If an IP address is included in the permitted segment, this address is permitted to log in to the MA5600T/MA5603T/MA5608T. In the preceding figure, the address of user A is permitted Telnet access.
- The denied IP address segment is an optional configuration item. If an IP address is included in the denied segment, this address is denied access to the MA5600T/MA5603T/MA5608T. In the preceding figure, the address of user B is denied Telnet access.

NOTE

The priority of the denied IP address segment is high. If an IP address is in both the permitted IP address segment and the denied IP address segment, the user with the IP address is not allowed to access the system.

26.5.3 Configuring a Firewall

A firewall monitors and determines whether data flows are allowed to enter an access device by analyzing data packets. It protects internal networks against unauthorized or unauthenticated access and attacks from external networks.

Context

The MA5600T/MA5603T/MA5608T filters data packets using the four firewall techniques listed in the following table.

Table 26-7 Firewall techniques supported by the MA5600T/MA5603T/MA5608T

Technique	Function	Feature
Firewall blacklist	A firewall blacklist filters data packets by source IP address.	Matching source IP addresses against a blacklist is simple, and packets can be quickly filtered. However, because data packets are filtered by only

Technique	Function	Feature
		one rule, this process lacks flexibility.
Firewall blacklist and advanced access control list (ACL) rules	The combination of a firewall blacklist and advanced ACL rules enables the MA5600T/MA5603T/MA5608T to further filter packets by advanced ACL rules.	Data packets are filtered based on a firewall blacklist and advanced ACL rules. The filter rules can be flexibly configured.
ACL-based packet filtering firewall	An ACL-based packet filtering firewall verifies data packets at the network layer and forwards or denies them according to the security policy.	Advantage: This technique supports more flexible configurations and better filtering capabilities than firewall blacklist. Disadvantages: • The packet filtering performance deteriorates sharply as the ACL complexity increases. • The MA5600T/MA5603T/MA5608T does not check the session status or analyze any data, and is vulnerable to IP spoofing attacks.
Unauthorized login prevention	The MA5600T/MA5603T/MA5608T prevents unauthorized logins by setting the IP address segments permitted by denied by the firewall for specified protocol types.	N/A

Procedure

- Configure a firewall blacklist.
  a. Run the **firewall blacklist item** command to add source IP addresses to the firewall blacklist.

     The data packet carrying a source IP address in the firewall blacklist is considered as untrustworthy.
  b. Run the **firewall blacklist enable** command to enable the firewall blacklist.
- Configure a combination of a firewall blacklist and advanced ACL rules.
  a. Run the **firewall blacklist item** command to add source IP addresses to the firewall blacklist.

b. Configure advanced ACL rules to filter out the data packets that carry a source IP addresses specified in the blacklist.
   i. Run the **acl** command to create an ACL. A firewall blacklist supports only an advanced ACL ranging from 3000 to 3999.
   ii. Run the **rule(adv acl)** command to create an advanced ACL rule.
   iii. Run the **quit** command to return to global config mode.

c. Run the **firewall blacklist enable acl-number** ***acl-number*** command to enable the firewall blacklist and apply the advanced ACL rule to the packets that carry a source IP address specified in the blacklist.

- Configure an ACL-based packet filtering firewall.
  a. Run the **acl** command to create an ACL. A firewall blacklist supports basic and advanced ACLs ranging from 2000 to 3999.
  b. Run the **rule** command to create an ACL rule.
     - Run the **rule(basic acl)** command to create a basic ACL rule.
     - Run the **rule(adv acl)** command to create an advanced ACL rule.
  c. Run the **quit** command to return to global config mode.
  d. To configure a firewall filtering rule for an METH port, run the **interface meth** command to enter METH mode; to configure a firewall filtering rule for a VLAN interface, run the **interface vlanif** command to enter VLAN interface mode.
  e. Run the **firewall packet-filter** command to apply the firewall filtering rule to the interface.

  NOTE
  When you run the **firewall packet-filter** command to activate an ACL, the MA5600T/MA5603T/MA5608T software determines the priority of ACL sub-rules. The ACL sub-rule configured earlier has a higher priority.

  f. Run the **firewall default** command to configure a packet filtering rule if a packet does not match any ACL rule.
  g. Run the **firewall enable** command to enable the firewall function for ACL-based packet filtering. The firewall is disabled by default.
- Configure a permitted or denied IP address segment to prevent unauthorized logins.
  a. Run the **sysman ip-access** command to configure an IP address segment that is permitted to connect to the MA5600T/MA5603T/MA5608T through Telnet, Secure Shell (SSH), or Simple Network Management Protocol (SNMP).
  b. Run the **sysman ip-refuse** command to configure an IP address segment that is denied to connect to the MA5600T/MA5603T/MA5608T through Telnet, SSH, or SNMP.
  c. Run the **sysman firewall** ***protocol-type*** **enable** command to enable the firewall function based on the protocol type (Telnet, SSH, or SNMP). The protocol-based firewall is disabled by default.

----End

Example

The following configurations are used as an example to configure a firewall:

- IP address added to the firewall blacklist: 192.168.10.18
- Aging time: 100 minutes

```
huawei(config)#firewall blacklist item 192.168.10.18 timeout 100
```

The following configurations are used as an example to configure a firewall:

- ACL type: advance
- Permitted IP address segment: 10.10.10.0
- Blacklist function: enabled

```
huawei(config)#acl 3000
huawei(config-acl-adv-3000)#rule permit ip source 10.10.10.0 0.0.0.255 destination
 10.10.10.20 0
huawei(config-acl-adv-3000)#quit
huawei(config)#firewall blacklist enable acl-number 3000
```

The following configurations are used as an example to configure a firewall to prevent some users from logging in to the maintenance network port on the MA5600T/MA5603T/MA5608T:

- Denied IP address segment: 172.16.25.0
- IP address of the maintenance network port: 172.16.25.28

```
huawei(config)#acl 3001
huawei(config-acl-adv-3001)#rule 5 deny icmp source 172.16.25.0 0.0.0.255 destin
ation 172.16.25.28 0
huawei(config-acl-adv-3001)#quit
huawei(config)#firewall enable
huawei(config)#interface meth 0
huawei(config-if-meth0)#firewall packet-filter 3001 inbound
 ACL applied successfully
```

The following configurations are used as an example to configure a firewall to allow some users to log in to the MA5600T/MA5603T/MA5608T:

- Protocol type: Telnet
- Permitted IP address segment: 10.10.5.1-10.10.5.254
- Login mode: Telnet

```
huawei(config)#sysman ip-access telnet 10.10.5.1 10.10.5.254
huawei(config)#sysman firewall telnet enable
```

The following configurations are used as an example to configure a firewall to allow some users to log in to the MA5600T/MA5603T/MA5608T:

- Protocol type: SSH
- Permitted IP address segment: 10.20.22.1-10.20.22.254
- Login mode: SSH

```
huawei(config)#sysman ip-access ssh 10.20.22.1 10.20.22.254
huawei(config)#sysman firewall ssh enable
```

The following configurations are used as an example to configure a firewall to refuse some users to log in to the MA5600T/MA5603T/MA5608T:

- Protocol type: SNMP
- Permitted IP address segment: 10.10.20.1-10.10.20.254
- Login mode: NMS

```
huawei(config)#sysman ip-refuse snmp 10.10.20.1 10.10.20.254
huawei(config)#sysman firewall snmp enable
```

27 Application Security

About This Chapter

This topic provides reference standards and protocols, and sub-features of the application security.

27.1 Introduction

User security refers to the security mechanism that ensures the security of access users, including the following features.

Feature	Description
HWTACACS	Similar to the RADIUS protocol, the MA5600T/MA5603T/MA5608T implements AAA functions for multiple users by communicating with the HWTACACS server in the client/server mode.
RAIO	Relay agent information option (RAIO) is the user physical location information provided by the device to the BRAS or DHCP server, such as the subrack ID, slot ID, and port ID on the device, when PITP and DHCP option 82 are enabled.
PITP	Policy information transfer protocol (PITP) is a protocol for implementing policy information transfer between the access device and the BRAS through Layer 2 P2P communication.
DHCP option 82	Add the user physical location information in the option 82 field of the DHCP request packet initiated by the user to co-work with the upper-layer authentication server to perform user authentication.
802.1x	802.1x is a port-based network access control protocol. Users connected to a port can access network resources through the port only after they pass the authentication.
Anti-IP spoofing	The system guards against the attack from users who forge IP addresses.

27.2 Relevant Standards and Protocols

PITP

TR101

802.1X

IEEE Std 802.1X-2001: Port-Based Network Access Control

RAIO

TR101

Anti IP Spoofing

None

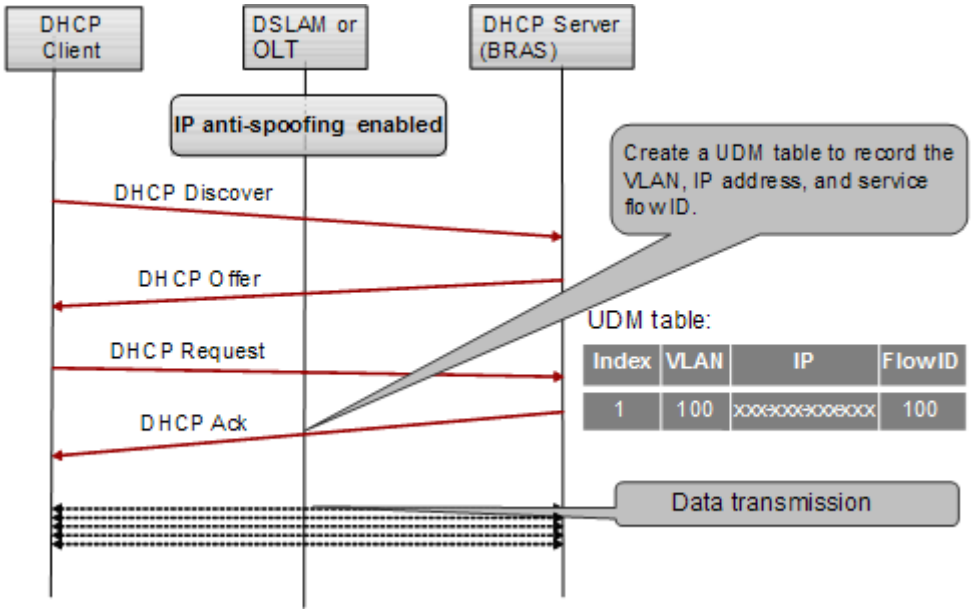
27.3 UDM

Principle

A unified device management (UDM) table is used to manage dialup user information. Dialup user information can be queried, managed, and filtered by creating and maintaining a UDM table. A UDM table contains the MAC address, IP address, VLAN, and flow ID lease period of dialup users. UDM tables are used for ARP proxy reply, MAC anti-spoofing, and IP anti-spoofing.

Generally, UDM tables are created when users go online. The following describes how a UDM table is generated when a user goes online through DHCP with IP anti-spoofing enabled.

The system monitors DHCP online and offline procedures of a user after it enables IP anti-spoofing. When the user goes online, the system dynamically obtains the IP address that has been allocated to the user, binds the IP address to the service flow, and records the IP address, VLAN, and service flow ID in the UDM table.



Similarly, the system creates a UDM table after it enables MAC anti-spoofing to record the MAC address, VLAN, and service flow ID. UDM tables are also used for ARP proxy reply and virtual MAC address (VMAC address).

The contents of a UDM table vary with the features for which the UDM table is used. For example, a UDM table contains IP address binding entries when IP anti-spoofing is enabled, and it contains MAC address binding entries when MAC anti-spoofing is enabled. Therefore, the capacity of a UDM table determines the actual specifications of these features. For example, if a UDM table supports a capacity of 8000 entries and 4000 entries are used for IP anti-spoofing, a maximum of 4000 entries can be used for MAC anti-spoofing.

Usage

The capacity of a UDM table depends on the product type and control board type. A dialup user (such as DHCP, PPPoE, DHCPv6, or SLAAC user) occupies an entry. The number of UDM entries supported by a service flow depends on the product type and service flow type.

27.4 AAA

AAA refers to authentication, authorization, and accounting. In the process that a user accesses network resources, through AAA, certain rights are authorized to the user if the user passes authentication, and the original data about the user accessing network resources is recorded.

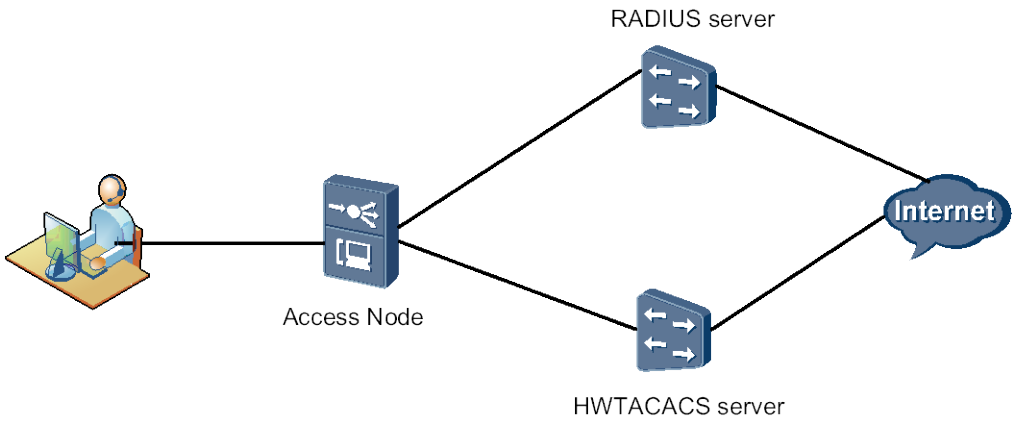
AAA includes three functions.

- Authentication: Checks whether a user is allowed to access network resources.
- Authorization: Determines what network resources a user can access.
- Accounting: Records the original data about the user accessing network resources.

AAA is generally applied to the users that access the Internet in the PPPoA, PPPoE, 802.1x, VLAN, WLAN, ISDN, or Admin SSH (associating the user name and the password with the domain name) mode.

In the existing network, 802.1x and Admin SSH correspond to the local AAA, that is, the access device functions as a local AAA server; PPPoE corresponds to the remote AAA, that is, the access device functions as the client of a remote AAA server.

The following picture shows an example network of the AAA application.



1. Authentication

   The MA5600T/MA5603T/MA5608T supports three authentication modes: non-authentication, local authentication, and remote authentication.

   - Non-authentication: The MA5600T/MA5603T/MA5608T trusts users and does not check the validity of the users. Generally, this mode is not adopted.
   - Local authentication: The user information (including the user name, password, and various attributes) is configured on the MA5600T/MA5603T/MA5608T, and the MA5600T/MA5603T/MA5608T authenticates the user. This authentication mode is fast and can reduce carrier's cost; however, the amount of information that can be stored is limited by the device hardware.
   - Remote authentication: The user information (including the user name, password, and various attributes of the user) is configured on an authentication server. The Remote Authentication Dial In User Service (RADIUS) protocol or HUAWEI Terminal Access Controller Access Control System (HWTACACS) protocol is used for remote authentication. The MA5600T/MA5603T/MA5608T serves as the authentication client and communicates with the RADIUS or HWTACACS server. When the RADIUS or HWTACACS server is faulty, the MA5600T/MA5603T/MA5608T can automatically switch to local authentication.

2. Authorization

   The MA5600T/MA5603T/MA5608T supports direct authorization, local authorization, HWTACACS authorization, and if-authenticated authorization.

   - Direct authorization: If trustful, a user can directly pass the authorization.
   - Local authorization: A user is locally authorized according to relevant attributes of the user configured on the MA5600T/MA5603T/MA5608T.
   - HWTACACS authorization: The HWTACACS server authorizes a user.
   - If-authenticated authorization: If a user passes the authentication and the authentication mode is not non-authentication, the user passes the authorization.

3. Accounting

   The MA5600T/MA5603T/MA5608T supports non-accounting and remote accounting.

   - Non-accounting: A user is not charged.

- Remote accounting: The MA5600T/MA5603T/MA5608T supports remote accounting through the AAA server.

The preceding figure shows that the AAA function can be implemented on the access device in the following three ways:

- The access device functions as a local AAA server. In this case, the local AAA needs to be configured. The local AAA does not support accounting.
- The access device functions as the client of a remote AAA server, and is connected to the RADIUS server through the RADIUS protocol, implementing the AAA. The RADIUS protocol, however, does not support authorization.
- The access device functions as the client of a remote AAA server, and is connected to the HWTACACS server through the HWTACACS protocol, implementing the AAA.

27.4.1 RADIUS

Definition

- Radius is short for the remote authentication dial-in user service. It is a distributed information interaction protocol with the client-server structure. Generally, it is used to manage a large number of distributed dial-in users.
- Radius implements the user accounting by managing a simple user authentication database.
- The authentication and accounting requests of users can be passed on to the Radius server through a network access server (NAS).

Principle

- When a user tries to access another network (or some network resources) by setting up a connection to the NAS through a network, the NAS forwards the user authentication and accounting information to the RADIUS server. The RADIUS protocol specifies the means of transmitting the user information and accounting information between the NAS and the RADIUS server.
- The RADIUS server receives the connection requests of users sent from the NAS, authenticates the user account and password contained in the user authentication data, and returns the required data to the NAS.

Message Flow of the RADIUS Protocol

The RADIUS server stores the user names and passwords in a unique user authentication database for authenticating the users. When a user wishes to connect to an NE through a device and then obtain the right to access the Internet or access certain network resources, the NE authenticates the user or the corresponding connection.

The NE sends the authentication, authorization, and accounting information of the user to the RADIUS server. The RADIUS protocol specifies how the NE and the RADIUS server should exchange the user information and the accounting information. The RADIUS server receives the connection request of the user, authenticates the user, and sends the necessary configuration information of the user to the NE. The exchange of authentication information between the NE and the RADIUS server is key protected. This protects the user password against any interception when the password is transmitted over an insecure network. Figure 27-1 shows the message flow between the RADIUS client and the RADIUS server.

Figure 27-1 Message flow between the RADIUS client and the RADIUS server

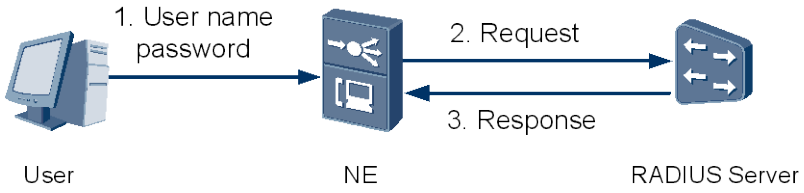
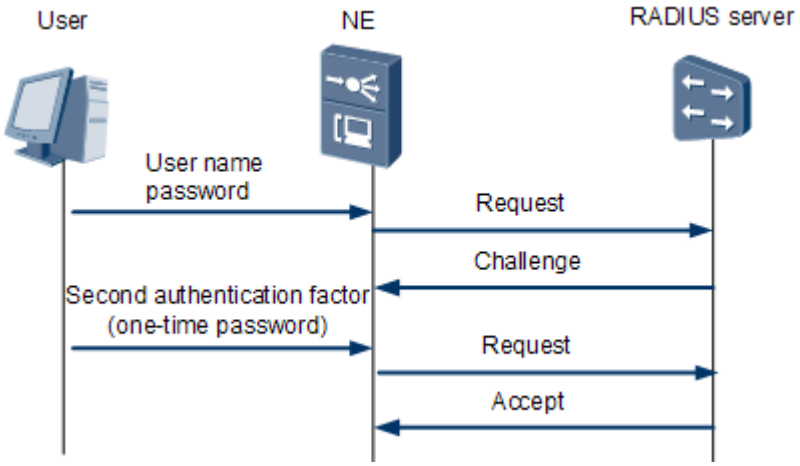


Figure 27-2 Two-factor authentication flow between the RADIUS client and the RADIUS server



NOTE

An NE refers to an access device that can function as a RADIUS client.

1. When a user logs in to the NE, the corresponding user name, password, and one-time password (required for two-factor authentication) are sent to the NE.
2. The RADIUS client on the NE receives the user name and password, and sends an authentication request to the RADIUS server.
3. The RADIUS server receives the legal request, authenticates the user, and sends the necessary authorization information of the user to the RADIUS client.

The authentication information exchanged between the RADIUS client and the RADIUS server must be encrypted before being transmitted over the network. Otherwise, the information may be intercepted when the network is insecure.

The accounting message flow is similar to the authentication/authorization message flow.

27.4.2 HWTACACS

Definition

HWTACACS (HUAWEI Terminal Access Controller Access Control System) is a security protocol enhanced based on TACACS+(draft-grant-tacacs-02). Similar to the RADIUS protocol, HWTACACS implements AAA functions for multiple users by communicating with the HWTACACS server in the client/server (C/S) mode.

Purpose

HWTACACS is used for the authentication, authorization, and accounting of access users and administrators.

Principle

Adopting the client/server architecture, HWTACACS is a protocol through which the network access server (NAS) (MA5600T/MA5603T/MA5608T) transmits the encrypted HWTACACS data packets to communicate with the HWTACACS database of the security server. The working mode is as follows:

- HWTACACS authentication. After being set up a connection with the NAS port, the NAS communicates with the daemon of the HWTACACS server. After the user input the user name, password, and dynamic password (if required) according to the prompt messages, the NAS sends these information to the daemon.
- HWTACACS authorization. After being authenticated, the user can be authorized. The NAS communicates with the daemon of the HWTACACS server, and then returns the accept or reject response of the authorization.

Differences Between HWTACACS and RADIUS

The HWTACACS message flow is similar to the RADIUS message flow. The difference is that, in the HWTACACS message flow, the server returns an authentication response rather than the user right after the user passes authentication. The user right is returned only when the authorization process is completed.

HWTACACS features more reliable transmission and encryption than RADIUS and is more suitable for security control. The following table lists the differences between HWTACACS and RADIUS.

HWTACACS	RADIUS
Uses TCP to ensure more reliable network transmission.	Uses UDP for transmission.
Encrypts the body of HWTACACS packets, except their header.	Encrypts only the password field of the authenticated packets.
Separated authorization and authentication.	Concurrent processing of authentication and authorization.
Applicable to security control.	Applicable to accounting.
Supports authorization of the configuration commands on the router.	Does not support the authorization of the configuration commands on the router.

27.4.3 Configuring the Local AAA

This topic describes how to configure the local AAA so that the user authentication can be performed locally.

Context

- The local AAA configuration is simple, which does not depend on the external server.
- The local AAA supports only authentication.

Procedure

Configure the AAA authentication scheme.

NOTE

- The authentication scheme specifies how all the users in an Internet service provider (ISP) domain are authenticated. The system supports up to 16 authentication schemes.
- The system has a default authentication scheme named **default**. It can be modified, but cannot be deleted.

1. Run the **aaa** command to enter the AAA mode.
2. Run the **authentication-scheme** command to add an authentication scheme.
3. Run the **authentication-mode local** command to configure the authentication mode of the authentication scheme.
4. Run the **quit** command to return to the AAA mode.

Step 1 Create a domain.

NOTE

- A domain is a group of users of the same type.
- In the user name format userid@domain-name (for example, huawei20041028@huawei.net), "userid" indicates the user name for authentication and "domain-name" followed by "@" indicates the domain name.
- The domain name for user login cannot exceed 15 characters, and the other domain names cannot exceed 20 characters.

1. In the AAA mode, run the **domain** command to create a domain.

Step 2 Refer the authentication scheme.

NOTE

You can refer an authentication scheme in a domain only after the authentication scheme is created.

1. In the domain mode, run the **authentication-scheme** command to reference the authentication scheme.
2. Run the **quit** command to return to the AAA mode.

Step 3 Configure a local user.

In the AAA mode, run the **local-user** *username* **service-type** command to create a local AAA user.

----End

Example

User1 in the isp domain adopts the local server for authentication. The authentication scheme is newscheme, the password is a123456, do as follows:

```
huawei(config)#aaa
huawei(config-aaa)#authentication-scheme newscheme
  Info: Create a new authentication scheme
```

```
huawei(config-aaa-authen-newscheme)#authentication-mode local
huawei(config-aaa-authen-newscheme)#quit
huawei(config-aaa)#domain isp
  Info: Create a new domain
huawei(config-aaa-domain-isp)#authentication-scheme newscheme
huawei(config-aaa-domain-isp)#quit
huawei(config-aaa)#local-user user1@isp service-type terminal password a123456
```

27.4.4 Configuring the Remote AAA (RADIUS Protocol)

The MA5600T/MA5603T/MA5608T is interconnected with the RADIUS server through the RADIUS protocol to implement authentication and accounting.

Context

- What is RADIUS:
  - Radius is short for the remote authentication dial-in user service. It is a distributed information interaction protocol with the client-server structure. Generally, it is used to manage a large number of distributed dial-in users.
  - Radius implements the user accounting by managing a simple user authentication database.
  - The authentication and accounting requests of users can be passed on to the Radius server through a network access server (NAS).
- Principle of RADIUS:
  - When a user tries to access another network (or some network resources) by setting up a connection to the NAS through a network, the NAS forwards the user authentication and accounting information to the RADIUS server. The RADIUS protocol specifies the means of transmitting the user information and accounting information between the NAS and the RADIUS server.
  - The RADIUS server receives the connection requests of users sent from the NAS, authenticates the user account and password contained in the user authentication data, and returns the required data to the NAS.
- Specification:
  - For the MA5600T/MA5603T/MA5608T, the RADIUS is configured based on each RADIUS server group.
  - In actual networking, a RADIUS server group can be any of the following:
    - An independent RADIUS server
    - A pair of primary/secondary RADIUS servers with the same configuration but different IP addresses
  - The following lists the attributes of a RADIUS server template:
    - IP addresses of primary and secondary servers
    - Shared key
    - RADIUS server type
- The configuration of the RADIUS protocol defines only the essential parameters for the information exchange between the MA5600T/MA5603T/MA5608T and the RADIUS server. To make the essential parameters take effect, the RADIUS server group should be referenced in a certain domain.

- The RADIUS attribute list defines the attribute parameters for interaction between the MA5600T/MA5603T/MA5608T and the RADIUS server. Table 27-1 describes the parameters.

Table 27-1 RADIUS attribute list

Parameter Code	Parameter Name	Description
1	User-Name	Indicates the user name for authentication.
2	Password	Indicates the user password for authentication. This parameter is valid only for PAP authentication.
3	Challenge-Password	Indicates the user password for authentication. This parameter is valid only for CHAP authentication.
4	NAS-IP-Address	Indicates the IP address of the access device. If the RADIUS server group is bound to an interface address, use the bound interface address; otherwise, use the address of the interface where packets are sent.
5	NAS-Port	Indicates the user access port. The format of this parameter is four-digit slot ID + two-digit card number + five-digit port number + 21-digit VLAN ID.
6	Service-Type	Indicates the user service type. The value of this parameter is 2 (frame) for access users and is 6 for remote management users. Currently, the MA5600T/MA5603T/MA5608T supports only 802.1x access users but not PPP, L2TP, or DHCP access users for RADIUS authentication.
7	Framed-Protocol	The value of this parameter is fixed to 1 (PPP) because ITU-T RFC 2856 does not define 802.1x for this parameter.
14	Login-IP-Host	Indicates the host IP address of a login user.
15	Login-Service	Indicates the login service type. The valid types are SSH, Rlogin, TCP Clear, PortMaster(proprietary), and

Parameter Code	Parameter Name	Description
		LAT.
24	State	If the access challenge packet that the RADIUS server sends to a device contains this parameter, the subsequent access request packet sent by the device to the RADIUS server must also contain this parameter of the same value as that is contained in the access challenge packet.
25	Class	If the access accept packet sent by the RADIUS server to a device contains this parameter, the subsequent charging request packet sent by the device to the RADIUS server must also contain this parameter of the same value. For a standard RADIUS server, a device can use the Class attribute to represent the CAR parameter.
27	Session-Timeout	Indicates the available remaining time in the unit of second. It is the user re-authentication time in the EAP challenge packet.
29	Termination-Action	Indicates the service termination mode. The valid modes are re-authentication and forcing users to go offline.
31	Calling-Station-Id	Allows the NAS to send the calling number.
32	NAS-Identifier	Indicates the host name of the device.
40	Acct-Status-Type	Indicates the charging packet type. • 1: charging start packet • 2: charging stop packet • 3: real-time charging packet
41	Acct-Delay-Time	Indicates the time for generating a charging packet in the unit of second.
42	Acct-Input-Octets	Indicates the number of upstream bytes in the unit of byte, kbyte, Mbyte, or Gbyte. The specific unit can be configured using commands.

Parameter Code	Parameter Name	Description
43	Acct-Output-Octets	Indicates the number of downstream bytes in the unit of byte, kbyte, Mbyte, or Gbyte. The specific unit can be configured using commands.
44	Acct-Session-Id	Indicates the charging connection number. The connection numbers for the charging start packet, real-time charging packet, and charging stop packet of the same connection must be the same.
45	Acct-Authentic	Indicates the user authentication mode. • 1: RADIUS authentication • 2: local authentication
46	Acct-Session-Time	Indicates the time for a user to go online in the unit of second.
47	Acct-Input-Packets	Indicates the number of upstream packets.
48	Acct-Output-Packets	Indicates the number of downstream packets.
49	Terminate-Cause	Indicates the user connection interruption cause. The valid values are as follows: • User-Request(1): The user actively goes offline. • Lost Carrier(2): The handshake fails, such as the EAPOL detection fails. • User Error(17): The user authentication fails or times out.
52	Acct-Input-Gigawords	Indicates the number of upstream bytes in the unit of 4Gbyte, kbyte, Mbyte, or Gbyte. The specific unit can be configured using commands.
53	Acct-Output-Gigawords	Indicates the number of downstream bytes in the unit of 4Gbyte, kbyte, Mbyte, or Gbyte. The specific unit can be configured using commands.
55	Event-Timestamp	Indicates the user online time in the unit of second. The value is the absolute number of seconds

Parameter Code	Parameter Name	Description
		counting from 1970-01-01 00:00:00.
60	CHAP-Challenge	Indicates the challenge field for CHAP authentication. This parameter is valid only for CHAP authentication.
61	NAS-Port-Type	Indicates the NAS port type.
79	EAP-Message	Carries EAP packets.
80	Message-Authenticator	Verifies validity of packets between the RADIUS server and RADIUS client to prevent malicious attacks.
85	Acct-Interim-Interval	Indicates the interval for real-time charging in the unit of second.
87	NAS-Port-Id	Indicates the user access port number. The format of this parameter uses the format when DHCP option 82 is in common raio mode.
88	Framed-Pool	Indicates the name and address segment number of the address pool. After being delivered by the RADIUS server, this parameter is filled to suboption 7 in user DHCP packets by the MA5600T/MA5603T/MA5608T.
26-29 **NOTE** The preceding parameters are RADIUS standard attributes. Starting from this row, the following parameters are Huawei-defined attributes.	Exec-Privilege	Indicates the priority of operation users such as SSH users. The value ranges from 0 to 15. • 0: common user • 1: operator • 2: administrator • 3-15: common user
26-60	Ip-Host-Address	Indicates the user IP address and MAC address that are contained in authentication and charging packets. The format is A.B.C.D HH:HH:HH:HH:HH:HH. The IP address and MAC address are separated by a space.
26-254	Version	Indicates the software version of the access device.
26-255	Product-ID	Indicates the product name.

NOTE

The super level user cannot be authenticated. You can query the user level by the command **display terminal user**.

Procedure

Configure the authentication scheme.

NOTE

- The authentication scheme specifies how all the users in an ISP domain are authenticated.
- The system supports up to 16 authentication schemes. The system has a default accounting scheme named **default**. It can only be modified, but cannot be deleted.

1. Run the **aaa** command to enter the AAA mode.
2. Run the **authentication-scheme** command to add an authentication scheme.
3. Run the **authentication-mode radius** command to configure the authentication mode of the authentication scheme.
4. Run the **quit** command to return to the AAA mode.

Step 1 Configure the accounting scheme.

NOTE

- The accounting scheme specifies how all the users in an ISP domain are charged.
- The system supports up to 128 accounting schemes. The system has a default accounting scheme named **default**. It can be modified, but cannot be deleted.

1. In the AAA mode, run the **accounting-scheme** command to add an AAA accounting scheme.
2. Run the **accounting-mode radius** command to configure the accounting mode.
3. Run the **accounting interim interval** command to set the interval of real-time accounting. By default, the interval is 0 minutes, that is, the real-time accounting is not performed.
4. Run the **quit** command to return to the AAA mode.

Step 2 Configure the RADIUS server template.

1. Run the **radius-server template** command to create an RADIUS server template and enter the RADIUS server template mode.
2. Run the **radius-server authentication** command to configure the IP address and the UDP port ID of the RADIUS server for authentication.

NOTE

- To guarantee normal communication between the MA5600T/MA5603T/MA5608T and the RADIUS server, before configuring the IP address and UDP port of the RADIUS server, make sure that the route between the RADIUS server and the MA5600T/MA5603T/MA5608T is in the normal state.
- Make sure that the configuration of the RADIUS service port of the MA5600T/MA5603T/MA5608T is consistent with the port configuration of the RADIUS server.

3. Run the **radius-server accounting** command to configure the IP address and the UDP port ID of the RADIUS server for accounting.
4. Run the **radius-server shared-key** command to configure the shared key of the RADIUS server.

NOTE

- The RADIUS client (MA5600T/MA5603T/MA5608T) and the RADIUS server use the MD5 algorithm to encrypt the RADIUS packets. They check the validity of the packets by setting the encryption key. They can receive the packets from each other and can respond to each other only when their keys are the same.
- By default, the shared key of the RADIUS server is **huawei**.

5. (Optional) Run the **radius-server timeout** command to set the response timeout time of the RADIUS server. By default, the timeout time is 5s.

   The MA5600T/MA5603T/MA5608T sends the request packets to the RADIUS server. If the RADIUS server does not respond within the response timeout time, the MA5600T/MA5603T/MA5608T re-transmits the request packets to the RADIUS to ensure that users can get corresponding services from the RADIUS server.

6. (Optional) Run the **radius-server retransmit** command to set the maximum re-transmit time of the RADIUS request packets. By default, the maximum re-transmit time is 3.

   When the re-transmit time of the RADIUS request packets to a RADIUS server exceeds the maximum re-transmit time, the MA5600T/MA5603T/MA5608T considers that its communication with the RADIUS server is interrupted, and therefore transmits the RADIUS request packets to another RADIUS server.

7. Run the **(undo)radius-server user-name domain-included** command to configure the user name (not) to carry the domain name when transmitted to the RADIUS server. By default, the user name of the RADIUS server carries the domain name.
   - An access user is named in the format of **userid@domain-name**, and the part after @ is the domain name. The MA5600T/MA5603T/MA5608T classifies a user into a domain according to the domain name.
   - If an RADIUS server group rejects the user name carrying the domain name, the RADIUS server group cannot be set or used in two or more domains. Otherwise, when some access users in different domains have the same user name, the RADIUS server considers that these users are the same because the names transmitted to the server are the same.

8. Run the **quit** command to return to the global config mode.

Step 3 Create a domain.

A domain is a group of users of the same type.

In the user name format userid@domain-name (for example, huawei20041028@huawei.net), "userid" indicates the user name for authentication and "domain-name" followed by "@" indicates the domain name.

The domain name for user login cannot exceed 15 characters, and the other domain names cannot exceed 20 characters.

1. Run the **aaa** command to enter the AAA mode.
2. In the AAA mode, run the **domain** command to create a domain.

Step 4 Use the authentication scheme.

You can use an authentication scheme in a domain only after the authentication scheme is created.

In the domain mode, run the **authentication-scheme** command to use the authentication scheme.

Step 5 Use the accounting scheme.

You can use an accounting scheme in a domain only after the accounting scheme is created.

In the domain mode, run the **accounting-scheme** command to use the accounting scheme.

Step 6 Use the RADIUS server template.

NOTE
You can use a RADIUS server template in a domain only after the RADIUS server template is created.

1. In the domain mode, run the **radius-server template** command to use the RADIUS server template.
2. Run the **quit** command to return to the AAA mode.

----End

Example

User1 in the isp domain adopts the HWTACACS protocol for authentication and accounting. The accounting interval is 10 minutes, the authentication password is a123456, HWTACACS server 10.10.66.66 functions as the primary authentication and accounting server, and HWTACACS server 10.10.66.67 functions as the standby authentication and accounting server. On the HWTACACS server, the authentication port ID is 1812, accounting port ID 1813, and other parameters adopt the default values. To perform the preceding configuration, do as follows:

```
huawei(config)#aaa
huawei(config-aaa)#authentication-scheme newscheme
huawei(config-aaa-authen-newscheme)#authentication-mode radius
huawei(config-aaa-authen-newscheme)#quit
huawei(config-aaa)#accounting-scheme newscheme
huawei(config-aaa-accounting-newscheme)#accounting-mode radius
huawei(config-aaa-accounting-newscheme)#accounting interim interval 10
huawei(config-aaa-accounting-newscheme)#quit
huawei(config)#radius-server template hwtest
huawei(config-radius-hwtest)#radius-server authentication 10.10.66.66 1812
huawei(config-radius-hwtest)#radius-server authentication 10.10.66.67 1812 secondary
huawei(config-radius-hwtest)#radius-server accounting 10.10.66.66 1813
huawei(config-radius-hwtest)#radius-server accounting 10.10.66.67 1813 secondary
huawei(config-radius-hwtest)#quit
huawei(config)#aaa
huawei(config-aaa)#domain isp
huawei(config-aaa-domain-isp)#authentication-scheme newscheme
huawei(config-aaa-domain-isp)#accounting-scheme newscheme
huawei(config-aaa-domain-isp)#radius-server hwtest
huawei(config-aaa-domain-isp)#quit
```

27.4.5 Configuration Example of the RADIUS Authentication and Accounting

The MA5600T/MA5603T/MA5608T is interconnected with the RADIUS server through the RADIUS protocol to implement authentication and accounting.

Service Requirements

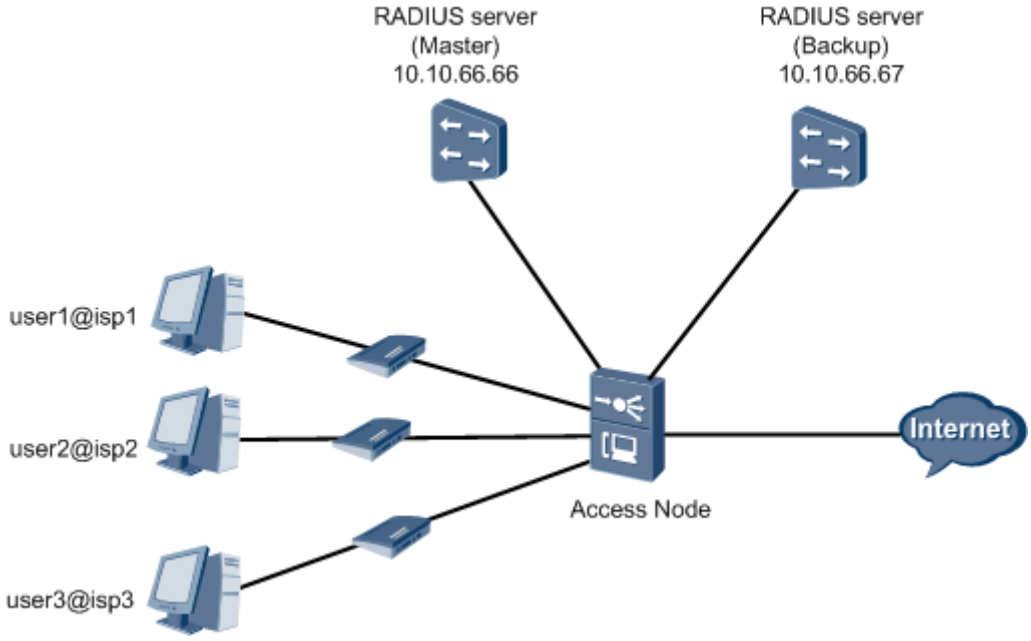
- The RADIUS server performs authentication and accounting for users in the ISP1 and ISP2 domains.

- The RADIUS server with the IP address 10.10.66.66 functions as the primary server for authentication and accounting.
- The RADIUS server with the IP address 10.10.66.67 functions as the secondary server for authentication and accounting.
- The authentication port number is 1812, and the accounting port number is 1813.
- Other parameters adopt the default settings.

Networking

Figure 27-3 shows an example network of the RADIUS Authentication and Accounting application.

Figure 27-3 Example network of the RADIUS Authentication and Accounting application.



Procedure

Configure the authentication scheme.

Configure authentication scheme named **newscheme** (users are authenticated through RADIUS).

```
huawei(config)#aaa
huawei(config-aaa)#authentication-scheme newscheme
  Info: Create a new authentication scheme
huawei(config-aaa-authen-newscheme)#authentication-mode radius
huawei(config-aaa-authen-newscheme)#quit
```

Step 1 Configure the accounting scheme.

Configure accounting scheme named **newscheme** (users are authenticated through RADIUS). the interval is 10 minutes.

```
huawei(config-aaa)#accounting-scheme newscheme
  Info: Create a new accounting scheme
huawei(config-aaa-accounting-newscheme)#accounting-mode radius
huawei(config-aaa-accounting-newscheme)#accounting interim interval 10
huawei(config-aaa-accounting-newscheme)#quit
huawei(config-aaa)#quit
```

Step 2 Configure the RADIUS protocol.

Create RADIUS server template named **hwtest** with the RADIUS server 10.10.66.66 as the primary authentication and accounting server, and the RADIUS server 10.10.66.67 as the secondary authentication and accounting server.

```
huawei(config)#radius-server template hwtest
 Note: Create a new server template
huawei(config-radius-hwtest)#radius-server authentication 10.10.66.66 1812
huawei(config-radius-hwtest)#radius-server authentication 10.10.66.67 1812 secondary
huawei(config-radius-hwtest)#radius-server accounting 10.10.66.66 1813
huawei(config-radius-hwtest)#radius-server accounting 10.10.66.67 1813 secondary
huawei(config-radius-hwtest)#quit
```

Step 3 Create a domain.

Create a domain named isp1.

```
huawei(config)#aaa
huawei(config-aaa)#domain isp1
  Info: Create a new domain
```

Step 4 Use the authentication scheme.

You can use an authentication scheme in a domain only after the authentication scheme is created.

```
huawei(config-aaa-domain-isp1)#authentication-scheme newscheme
```

Step 5 Use the accounting scheme.

You can use an accounting scheme in a domain only after the accounting scheme is created.

```
huawei(config-aaa-domain-isp1)#accounting-scheme newscheme
```

Step 6 Use the RADIUS server template.

You can use a RADIUS server template in a domain only after the RADIUS server template is created.

```
huawei(config-aaa-domain-isp1)#radius-server hwtest
huawei(config-aaa-domain-isp1)#quit
```

----End

Result

User 1 in ISP 1 can pass authentication only if both the user name and password are correct, and then can log in to the MA5600T/MA5603T/MA5608T. Then, the user starts to be accounted.

Configuration File

```
aaa
authentication-scheme newscheme
authentication-mode radius
quit
accounting-scheme newscheme
accounting-mode radius
accounting interim interval 10
quit
quit
radius-server template hwtest
radius-server authentication 10.10.66.66 1812
radius-server authentication 10.10.66.67 1812 secondary
radius-server accounting 10.10.66.66 1813
radius-server accounting 10.10.66.67 1813 secondary
quit
aaa

domain isp1
authentication-scheme newscheme
accounting-scheme newscheme
radius-server hwtest
quit
```

27.4.6 Configuring the Remote AAA (HWTACACS Protocol)

The MA5600T/MA5603T/MA5608T is interconnected with the HWTACACS server through the HWTACACS protocol to implement authentication, authorization, and accounting.

Context

- What is HWTACACS:
  - HWTACACS is a security protocol with enhanced functions on the base of TACACS+(draft-grant-tacacs-02). Similar to the RADIUS protocol, HWTACACS implements multiple subscriber AAA functions through communications with the HWTACACS server in the client/server (C/S) mode.
  - HWTACACS is used for the authentication, authorization, and accounting for the 802.1 access users and management users.
- Principle of HWTACACS:

  Adopting the client/server architecture, HWTACACS is a protocol through which the NAS (MA5600T/MA5603T/MA5608T) transmits the encrypted HWTACACS data packets to communicate with the HWTACACS database of the security server. The working mode is as follows:
  - HWTACACS authentication. When the remote user connects to the corresponding port of the NAS, the NAS communicates with the daemon of the HWTACACS server, and obtains the prompt of entering the user name from the daemon. Then, the NAS displays the message to the user. When the remote user enters the user name, the NAS transmits the user name to the daemon. Then, the NAS obtains the prompt of entering the password, and displays the message to the user. After the remote user enters the password, the NAS transmits the password to the daemon.

- HWTACACS authorization. After being authenticated, the user can be authorized. The NAS communicates with the daemon of the HWTACACS server, and then returns the accept or reject response of the authorization.

NOTE

- The HWTACACS configuration only defines the parameters used for data exchange between the MA5600T/MA5603T/MA5608T and the HWTACACS server. To make these parameters take effect, you need to use the HWTACACS server group in a domain.
- The settings of an HWTACACS server template can be modified regardless of whether the template is bound to a server or not.

Procedure

Configure the AAA authentication scheme.

The authentication scheme specifies how all the users in an ISP domain are authenticated.

The system supports up to 16 authentication schemes. The system has a default authentication scheme named **default**. It can be modified, but cannot be deleted.

1. Run the **aaa** command to enter the AAA mode.
2. Run the **authentication-scheme** command to add an authentication scheme.
3. Run the **authentication-mode local** command to configure the authentication mode of the authentication scheme. Use the HWTACACS protocol to authenticate users.
4. Run the **quit** command to return to the AAA mode.

Step 1 Configure the AAA authorization scheme.

The authorization scheme specifies how all the users in an ISP domain are authorized.

1. In the AAA mode, run the **authorization-scheme** command to add an AAA authorization scheme.
2. Run the **authorization-mode hwtacacs** command to configure the authorization mode.
3. Run the **quit** command to return to the AAA mode.
4. Run the **quit** command to return to the global config mode.

Step 2 Configure the AAA accounting scheme.

The accounting scheme specifies how all the users in an ISP domain are charged.

The system supports up to 128 accounting schemes. The system has a default accounting scheme named **default**. It can be modified, but cannot be deleted.

1. In the AAA mode, run the **accounting-scheme** command to add an AAA accounting scheme.
2. Run the **accounting-mode hwtacacs** command to configure the accounting mode. By default, the accounting is not performed.
3. Run the **accounting interim interval** command to set the interval of real-time accounting. By default, the interval is 0 minutes, that is, the real-time accounting is not performed.
4. Run the **quit** command to return to the AAA mode.

Step 3 Configure the HWTACACS protocol.

The configuration of the HWTACACS protocol of the MA5600T/MA5603T/MA5608T is on the basis of the HWTACACS server group. In actual networking scenarios, an HWTACACS server group can be an independent HWTACACS server or a combination of two

HWTACACS servers, that is, a primary server and a secondary server with the same configuration but different IP addresses.

Each HWTACACS server template contains the primary/secondary server IP address, shared key, and HWTACACS server type.

Primary and secondary authentication, accounting, and authorization servers can be configured. The IP address of the primary server, however, must be different from that of the secondary server. Otherwise, the configuration of primary and secondary servers will fail. By default, the IP addresses of the primary and secondary servers are both 0.0.0.0.

1. Run the **hwtacacs-server template** command to create an HWTACACS server template and enter the HWTACACS server template mode.
2. Run the **hwtacacs-server authentication** command to configure a primary authentication server. You can select **secondary** to configure a secondary authentication server.

NOTE

- To ensure normal communication between the MA5600T/MA5603T/MA5608T and the HWTACACS server, before configuring the IP address and the UDP port of the HWTACACS server, make sure that the route between the HWTACACS server and the MA5600T/MA5603T/MA5608T is in the normal state.
- Make sure that the HWTACACS server port of the MA5600T/MA5603T/MA5608T is the same as the port of the HWTACACS server.

3. Run the **hwtacacs-server accounting** command to configure a primary accounting server. You can select **secondary** to configure a secondary accounting server.
4. Run the **hwtacacs-server authorization** command to configure a primary authorization server. You can select **secondary** to configure a secondary authorization server.
5. (Optional) Run the **hwtacacs-server shared-key** command to configure the shared key of the HWTACACS server.

NOTE

- The HWTACACS client (MA5600T/MA5603T/MA5608T) and the HWTACACS server use the MD5 algorithm to encrypt the HWTACACS packets. They check the validity of the packets by configuring the encryption key. They can receive the packets from each other and can respond to each other only when their keys are the same.
- By default, the HWTACACS server does not have a key.

6. (Optional) Run the **hwtacacs-server timer response-timeout** to set the response timeout time of the HWTACACS server.

NOTE

- If the HWTACACS server does not respond to the HWTACACS request packets within the timeout time, the communication between the MA5600T/MA5603T/MA5608T and the current HWTACACS server is considered as interrupted.
- By default, the response timeout time of the HWTACACS server is 5s.

7. (Optional) In the global config mode, run the **hwtacacs-server accounting-stop-packet** command to configure the re-transmission mechanism of the accounting-stop packets of the HWTACACS server.

NOTE

- To prevent the loss of the accounting packets, the MA5600T/MA5603T/MA5608T supports the re-transmission of the accounting-stop packets of the HWTACACS server.
- By default, the re-transmit time of the accounting-stop packets of the HWTACACS server is 100.

8. (Optional) Run the **(undo)hwtacacs-server user-name domain-included** command to configure the user name (not) to carry the domain name when transmitted to the HWTACACS server.
   - By default, the user name of the HWTACACS server carries the domain name.

- After the **undo hwtacacs-server user-name domain-included** command is executed, the domain name is deleted from the user name when the client sends authentication and authorization requests to the HWTACACS server. The domain name in the user name of the accounting request is, however, reserved. This is to ensure that the users can be distinguished from each other in the accounting.

9. Run the **quit** command to return to the global config mode.

Step 4 Create a domain.

A domain is a group of users of the same type.

In the user name format userid@domain-name (for example, huawei20041028@huawei.net), "userid" indicates the user name for authentication and "domain-name" followed by "@" indicates the domain name.

The domain name for user login cannot exceed 15 characters, and the other domain names cannot exceed 20 characters.

1. Run the **aaa** command to enter the AAA mode.
2. In the AAA mode, run the **domain** command to create a domain.

Step 5 Use the authentication scheme.

You can use an authentication scheme in a domain only after the authentication scheme is created.

In the domain mode, run the **authentication-scheme** command to use the authentication scheme.

Step 6 Use the accounting scheme.

You can use an accounting scheme in a domain only after the accounting scheme is created.

In the domain mode, run the **accounting-scheme** command to use the accounting scheme.

Step 7 Use the authorization scheme.

You can use an authorization scheme in a domain only after the authorization scheme is created.

In the domain mode, run the **authorization-mode** command to use the authorization scheme.

Step 8 Use the HWTACACS server template.

You can use an HWTACACS server template in a domain only after the HWTACACS server template is created.

1. In the domain mode, run the **hwtacacs-server** command to use the HWTACACS server template.
2. Run the **quit** command to return to the AAA mode.

----End

Example

User1 in the isp domain adopts the HWTACACS protocol for authentication, authorization, and accounting. The accounting interval is 10 minutes, the authentication password is a123456, HWTACACS server 10.10.66.66 functions as the primary authentication, authorization, and accounting server, and HWTACACS server 10.10.66.67 functions as the

standby authentication, authorization, and accounting server. On the HWTACACS server, the parameters adopt the default values. To perform the preceding configuration, do as follows:

```
huawei(config)#aaa
huawei(config-aaa)#authentication-scheme newscheme
huawei(config-aaa-authen-newscheme)#authentication-mode hwtacacs
huawei(config-aaa-authen-newscheme)#quit
huawei(config-aaa)#authorization-scheme newscheme
huawei(config-aaa-author-newscheme)#authorization-mode hwtacacs
huawei(config-aaa-author-newscheme)#quit
huawei(config-aaa)#accounting-scheme newscheme
huawei(config-aaa-accounting-newscheme)#accounting-mode hwtacacs
huawei(config-aaa-accounting-newscheme)#accounting interim interval 10
huawei(config-aaa-accounting-newscheme)#quit
huawei(config)#hwtacacs-server template hwtest
huawei(config-hwtacacs-hwtest)#hwtacacs-server authentication 10.10.66.66
huawei(config-hwtacacs-hwtest)#hwtacacs-server authentication 10.10.66.67 secondary
huawei(config-hwtacacs-hwtest)#hwtacacs-server authorization 10.10.66.66
huawei(config-hwtacacs-hwtest)#hwtacacs-server authorization 10.10.66.67 secondary
huawei(config-hwtacacs-hwtest)#hwtacacs-server accounting 10.10.66.66
huawei(config-hwtacacs-hwtest)#hwtacacs-server accounting 10.10.66.67 secondary
huawei(config-hwtacacs-hwtest)#quit
huawei(config)#aaa
huawei(config-aaa)#domain isp
huawei(config-aaa-domain-isp)#authentication-scheme newscheme
huawei(config-aaa-domain-isp)#authorization-scheme newscheme
huawei(config-aaa-domain-isp)#accounting-scheme newscheme
huawei(config-aaa-domain-isp)#hwtacacs-server hwtest
huawei(config-aaa-domain-isp)#quit
```

27.4.7 Configuration Example of the HWTACACS Authentication (802.1X access user)

The MA5600T/MA5603T/MA5608T is interconnected with the HWTACACS server through the HWTACACS protocol to implement authentication, authorization, and accounting.

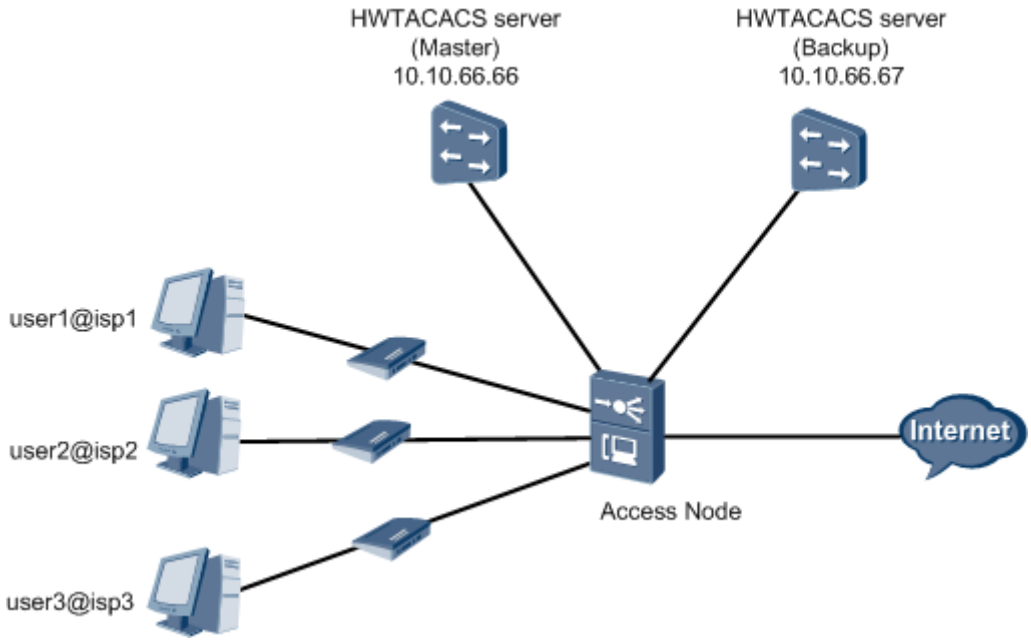
Service Requirements

- The HWTACACS server performs authentication, authorization, and accounting for 802.1X access users.
- The user logs in to the server carrying the domain name.
- The HWTACACS server with the IP address 10.10.66.66 functions as the primary server for authentication, authorization, and accounting.
- The HWTACACS server with the IP address 10.10.66.67 functions as the secondary server for authentication, authorization, and accounting.
- Other parameters adopt the default settings.

Networking

Figure 27-4 shows an example network of the HWTACACS authentication.

Figure 27-4 Example network of the HWTACACS authentication



Procedure

Configure an authentication scheme.

Configure authentication scheme named **newscheme** (users are authenticated through HWTACACS).

```
huawei(config)#aaa
huawei(config-aaa)#authentication-scheme newscheme
huawei(config-aaa-authen-newscheme)#authentication-mode hwtacacs
huawei(config-aaa-authen-newscheme)#quit
```

Step 1 Configure an authorization scheme.

Configure authorization scheme named **newscheme** (users are authorized through HWTACACS).

```
huawei(config-aaa)#authorization-scheme newscheme
huawei(config-aaa-author-newscheme)#authorization-mode hwtacacs
huawei(config-aaa-author-newscheme)#quit
```

Step 2 Configure the accounting scheme.

Configure accounting scheme named **newscheme** (users are authenticated through HWTACACS). the interval is 10 minutes.

```
huawei(config-aaa)#accounting-scheme newscheme
huawei(config-aaa-accounting-newscheme)#accounting-mode hwtacacs
huawei(config-aaa-accounting-newscheme)#accounting interim interval 10
huawei(config-aaa-accounting-newscheme)#quit
huawei(config-aaa)#quit
```

Step 3 Configure the HWTACACS protocol.

Create HWTACACS server template named **hwtest** with the HWTACACS server 10.10.66.66 as the primary authentication, authorization and accounting server, and the HWTACACS server 10.10.66.67 as the secondary authentication, authorization and accounting server.

```
huawei(config)#hwtacacs-server template hwtest
  Create a new HWTACACS-server template
huawei(config-hwtacacs-radtest)#hwtacacs-server authentication 10.10.66.66
huawei(config-hwtacacs-radtest)#hwtacacs-server authentication 10.10.66.67 secondary
huawei(config-hwtacacs-hwtest)#hwtacacs-server authorization 10.10.66.66
huawei(config-hwtacacs-hwtest)#hwtacacs-server authorization 10.10.66.67 secondary
huawei(config-hwtacacs-radtest)#hwtacacs-server accounting 10.10.66.66
huawei(config-hwtacacs-radtest)#hwtacacs-server accounting 10.10.66.67 secondary
huawei(config-hwtacacs-radtest)#quit
```

Step 4 Configure the 802.1X authentication.

1. Enable the 802.1X global switch. Enable the 802.1X authentication for ports 1, 2, and 3. The 802.1X needs to be triggered by DHCP. Therefore, the DHCP-trigger authentication must be enabled.

   ```
   huawei(config)#dot1x enable
   huawei(config)#dot1x service-port 1
   huawei(config)#dot1x service-port 2
   huawei(config)#dot1x service-port 3
   huawei(config)#dot1x dhcp-trigger enable
   ```

2. Configure an 802.1X parameters. In the local termination authentication, the 802.1X parameters should be configured to be in the EAP termination mode. The count of allowed handshake failure is 1 and the handshake interval is 20s.

   ```
   huawei(config)#dot1x keepalive retransmit 1 interval 20 service-port 1
   huawei(config)#dot1x keepalive retransmit 1 interval 20 service-port 2
   huawei(config)#dot1x keepalive retransmit 1 interval 20 service-port 3
   huawei(config)#dot1x eap-end service-port 1
   huawei(config)#dot1x eap-end service-port 2
   huawei(config)#dot1x eap-end service-port 3
   ```

Step 5 Create a domain.

Create a domain named isp1.

```
huawei(config)#aaa
huawei(config-aaa)#domain isp1
  Info: Create a new domain
```

Step 6 Use the authentication scheme.

You can use an authentication scheme in a domain only after the authentication scheme is created.

```
huawei(config-aaa-domain-isp1)#authentication-scheme newscheme
```

Step 7 Use the authorization scheme.

You can use an authorization scheme in a domain only after the authorization scheme is created.

```
huawei(config-aaa-domain-isp1)#authorization-scheme newscheme
```

Step 8 Use the accounting scheme.

You can use an accounting scheme in a domain only after the accounting scheme is created.

```
huawei(config-aaa-domain-isp1)#accounting-scheme newscheme
```

Step 9 Bind the HWTACACS server template.

You can use an HWTACACS server template in a domain only after the HWTACACS server template is created.

```
huawei(config-aaa-domain-isp1)#hwtacacs-server hwtest
```

----End

Result

User 1 in ISP 1 can pass authentication only if both the user name and password are correct, and then can log in to the MA5600T/MA5603T/MA5608T. Then, the user starts to be accounted.

Configuration File

```
aaa
authentication-scheme newscheme
authentication-mode hwtacacs
quit
authorization-scheme newscheme
authorization-mode hwtacacs
quit
accounting-scheme newscheme
accounting-mode hwtacacs
accounting interim interval 10
quit
quit
hwtacacs-server template hwtest
hwtacacs-server authentication 10.10.66.66
hwtacacs-server authentication 10.10.66.67 secondary
hwtacacs-server authorization 10.10.66.66
hwtacacs-server authorization 10.10.66.67 secondary
hwtacacs-server accounting 10.10.66.66
hwtacacs-server accounting 10.10.66.67 secondary
quit
dot1x enable
dot1x service-port 1
dot1x service-port 2
dot1x service-port 3
dot1x dhcp-trigger enable
dot1x keepalive retransmit 1 interval 20 service-port 1
dot1x keepalive retransmit 1 interval 20 service-port 2
dot1x keepalive retransmit 1 interval 20 service-port 3
dot1x eap-end service-port 1
dot1x eap-end service-port 2
dot1x eap-end service-port 3
```

```
domain isp1
authentication-scheme newscheme
authorization-scheme newscheme
accounting-scheme newscheme
hwtacacs-server hwtest
```

27.4.8 Configuration Example of HWTACACS Authentication (Management User)

The MA5600T/MA5603T/MA5608T allows the management user of the device to log in to the system by the HWTACACS authentication mode.

Prerequisites

- The route from the MA5600T/MA5603T/MA5608T to the HWTACACS server must be configured.
- The management user information (user name@domain and password) must be configured on the HWTACACS server.

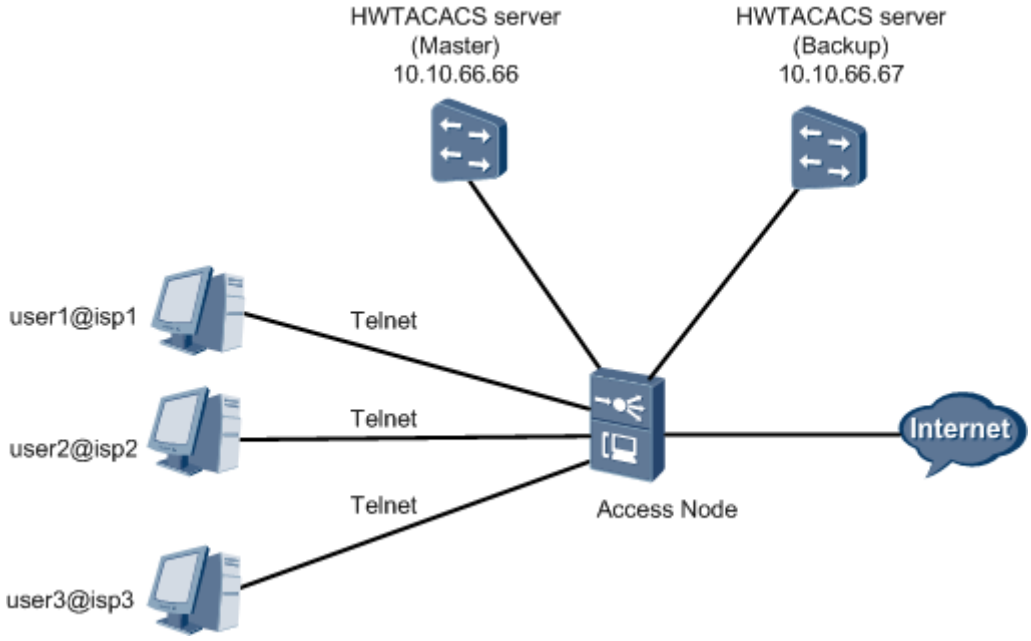
Service Requirements

- The HWTACACS server performs authentication for management user of domain **isp1**.
- The user logs in to the server carrying the domain name.
- The HWTACACS server with the IP address 10.10.66.66 functions as the primary server for authentication.
- The HWTACACS server with the IP address 10.10.66.67 functions as the secondary server for authentication.
- Other parameters adopt the default settings.

Networking

Figure 27-5 shows an example network of HWTACACS authentication.

Figure 27-5 Example network of HWTACACS authentication



Procedure

Configure the authentication scheme.

Configure authentication scheme named **login-auth** (users are authenticated through HWTACACS).

```
huawei(config)#aaa
huawei(config-aaa)#authentication-scheme login-auth
huawei(config-aaa-authen-login-auth)#authentication-mode hwtacacs
huawei(config-aaa-authen-login-auth)#quit
```

Step 1 Configure the HWTACACS protocol.

Create HWTACACS server template named **ma56t-login** with HWTACACS server 10.10.66.66 as the primary authentication server, and HWTACACS server 10.10.66.67 as the secondary authentication server.

```
huawei(config)#hwtacacs-server template ma56t-login
  Create a new HWTACACS-server template
huawei(config-hwtacacs-ma56t-login)#hwtacacs-server authentication 10.10.66.66 1812
huawei(config-hwtacacs-ma56t-login)#hwtacacs-server authentication 10.10.66.67 1812
secondary
huawei(config-hwtacacs-ma56t-login)#quit
```

Step 2 Create a domain named **isp1**.

NOTE

- A domain is a group of users of the same type.
- In the user name format userid@domain-name (for example, huawei20041028@huawei.net), "userid" indicates the user name for authentication and "domain-name" followed by "@" indicates the domain name.

- The domain name for user login cannot exceed 15 characters, and the other domain names cannot exceed 20 characters.

```
huawei(config)#aaa
huawei(config-aaa)#domain isp1
  Info: Create a new domain
```

Step 3 Use the authentication scheme **login-auth**.

You can use an authentication scheme in a domain only after the authentication scheme is created.

```
huawei(config-aaa-domain-isp1)#authentication-scheme login-auth
```

Step 4 Bind the HWTACACS server template **ma56t-login** to the user.

You can use an HWTACACS server template in a domain only after the HWTACACS server template is created.

```
huawei(config-aaa-domain-isp1)#hwtacacs-server ma56t-login
```

----End

Result

- When the HWTACACS server is reachable, the management user can log in to the MA5600T/MA5603T/MA5608T through SSH. After entering the user name and password specified on the HWTACACS server, the management user can successfully log in to the MA5600T/MA5603T/MA5608T.
- When the HWTACACS server is unreachable, the management user cannot log in to the MA5600T/MA5603T/MA5608T through SSH by entering the user name and password specified on the HWTACACS server.

Configuration File

```
huawei(config)#aaa
huawei(config-aaa)#authentication-scheme login-auth
huawei(config-aaa-authen-login-auth)#authentication-mode hwtacacs
huawei(config-aaa-authen-login-auth)#quit
huawei(config-aaa)#quit
huawei(config)#hwtacacs-server template ma56t-login
huawei(config-hwtacacs-ma56t-login)#hwtacacs-server authentication 10.10.66.66 1812
huawei(config-hwtacacs-ma56t-login)#hwtacacs-server authentication 10.10.66.67 1812
secondary
huawei(config-hwtacacs-ma56t-login)#quit
huawei(config)#aaa
huawei(config-aaa)#domain isp1
huawei(config-aaa-domain-isp1)#authentication-scheme login-auth
huawei(config-aaa-domain-isp1)#hwtacacs-server ma56t-login
huawei(config-aaa-domain-isp1)#quit
huawei(config-aaa)#quit
```

27.5 802.1X

IEEE 802.1X (hereinafter referred to as 802.1X) is a port-based network access control protocol.

27.5.1 Introduction

Definition

IEEE 802.1X (hereinafter referred to as 802.1X) is a port-based network access control protocol.

If a user connected to a port can pass the authentication, the user can access the resources in the network. In case of a failure to pass the authentication, the user cannot access the resources in the network. That is, the physical connection is cut off.

The 802.1X port can be a physical port or a logical port.

Purpose

The MA5600T/MA5603T/MA5608T supports the port-based access authentication mode as specified in the standard. In addition, it extends and optimizes this authentication mode. As a result, the system security is improved and the system management function is enhanced.

27.5.2 Principle

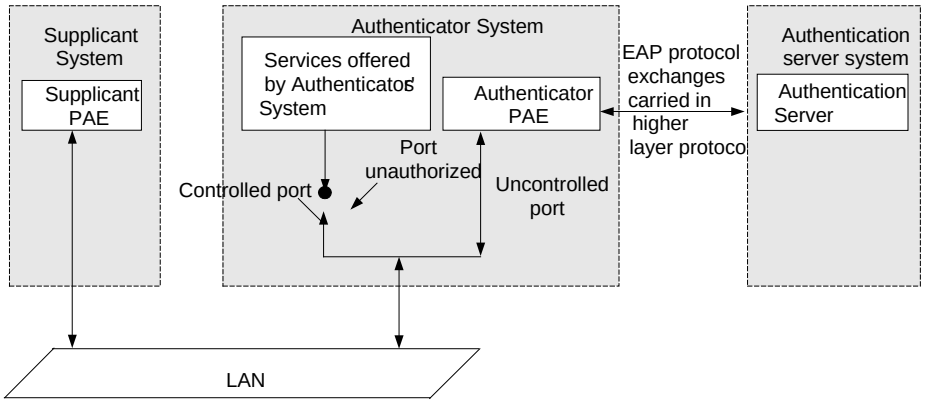
Protocol System

802.1X defines the port-based network access control from the following aspects:

- The access device provides the authentication control function of the access port (physical port or logical port).
- Before a port passes the authentication, the port is disabled and the users connected to the port cannot access the network resources.
- If the port passes the authentication, the port is enabled and the users can access the network. If the port does not pass the authentication, the port is disabled and the users cannot access the network.

The 802.1X system defines three functional entities: supplicant system, authenticator system, and authentication server system. Figure 27-6 shows the 802.1X system architecture.

Figure 27-6 802.1X system architecture



In general, the digital user terminal provides the functions of the supplicant system entity and needs to be installed with the 802.1X client software, through which the supplicant system initiates authentication and quits authentication.

The authenticator system authenticates the request from the supplicant. An authenticator system is usually an 802.1X-enabled network device, providing a service port for the supplicant. The service port can be a physical port or a logical port, and implements the 802.1X authentication of access users.

The authentication server is an entity that provides the authentication service for the authenticator system. The 802.1X authentication server is usually located in the operator's AAA center.

The ports of the authenticator system can be controlled ports or uncontrolled ports.

- A controlled port is used to transmit the authenticated service packets. If a user passes the authentication, the controlled port changes to the authenticated state, and then the port can transmit the service packets. If the user fails to pass the authentication, the controlled port changes to the unauthenticated state, and the port cannot transmit the service packets.
- An uncontrolled port is always in the bi-directional connection state and can transmit authentication protocol packets, regardless of the authentication state (authenticated state or unauthenticated state) of the controlled port.

Feature Implementation

The device supports user access authentication based on port, service flow, or port + MAC address.

- In the case of the authentication based on port, the port state is in down before authentication. Once the authentication is passed, the port state is up and all the service flows of the port are enabled.
- In the case of the authentication based on service flow, a service flow is disabled before authentication. Once the authentication is passed, the service flow is enabled and in such a case, all user terminals of the service virtual port can access the network.
- In the case of the authentication based on port + MAC address, the port state is up and only the packets with the MAC address that passes authentication are allowed to pass through the port.

In the case of the authentication by service virtual port, a service virtual port can be any of the following:

- An xDSL ATM service virtual port which is identified by the PVC or the PVC plus the user VLAN
- An xDSL PTM service virtual port which is identified by the user VLAN

The MA5600T/MA5603T/MA5608T supports the 802.lX authentication triggered by EAPoL or DHCP packets. You can set the method for EAPoL or DHCP packets to trigger the 802.1X authentication according to the terminal capability.

With the 802.1X protocol running, the MA5600T/MA5603T/MA5608T works as an authenticator and receives the authentication requests from the users. In the case of a remote authentication, the MA5600T/MA5603T/MA5608T sends the authentication information to the RADIUS server for authentication. If an access port passes the authentication of the RADIUS server, it is enabled.

The MA5600T/MA5603T/MA5608T supports the EAP termination and EAP relay modes.

- In the EAP termination mode, the MA5600T/MA5603T/MA5608T abstracts the user authentication information from the EAP packets, encapsulates the information into the corresponding attribute of the RADIUS protocol, and then sends the information to the RADIUS server for authentication.
- In the EAP relay mode, the MA5600T/MA5603T/MA5608T encapsulates the EAP packets into the corresponding attribute of the RADIUS protocol, and sends the packets to the RADIUS server for authentication. In this mode, the RADIUS server needs to process the EAP packets.

27.6 Anti-IP Spoofing

The anti-IP spoofing function prevents a user from forging IPaddresses to initiate attacks so that network security is improved.

27.6.1 Introduction

Definition

IP spoofing is an attack in which malicious users send packets with forged IP addresses to attack the system. Malicious users can forge the IP addresses of authorized users to damage the services of these users.

Anti-IP spoofing is a countermeasure that is taken by the system to prevent a user from attacking the system with a forged IP address.

Purpose

To protect the system and the network of a carrier, for authorized users that access the network following the DHCP online process, the system dynamically binds MAC addresses and allows the users with trustful IP addresses to enter the network. Users with untrusted IP addresses are prohibited from entering the network.

For authorized users that do not access a carrier's network following the DHCP online process, the system binds the static IP addresses of users and allows the users with trustful IP addresses to enter the network.

Benefits

Benefits to carriers: Anti-IP spoofing, using dynamic or static IP address binding, protects the carrier's network from being attacked.

Benefits to users: Anti-IP spoofing, using dynamic or static IP address binding, enhances the security of user services.

27.6.2 Principle

Dynamic IP Address Binding for Anti-IP Spoofing

- After the dynamic IPv4 address learning function is disabled, the system monitors users' DHCP online and offline processes. When a user goes online, the system dynamically obtains the user's source IPv4 address and binds the user's source IPv4 address to a traffic stream.
- The system only allows the packets with source IPv4 addresses bound to the user port or traffic stream to pass through.
- When a user goes offline, the system unbinds the user's source IPv4 address from the traffic stream.

After anti-IP spoofing by binding dynamic IP addresses is enabled, the access device will modify the exchange identification (XID) of the DHCP packet sent by the user, so that the XID of the DHCP packet sent by the DHCP client is different from that of the DHCP packet received by the DHCP server. Generally, the DHCP server does not verify the XID, and therefore services are not affected. If the carrier adds information into the XID of the packet sent by the DHCP client for DHCP server verification (this is not defined in the standard), the verification may fail and services will be affected.

NOTE
XID is a field carried by the DHCP packet, and it is defined by the standard. The XID is equivalent to the serial number of the DHCP packet.

Static IP Address Binding for Anti-IP Spoofing

The access device allows you to bind IPv4 addresses to user ports. After IP addresses are bound to a user port, the user port only allows packets with IP addresses bound to the port to pass through. This improves system security.

Recovery of IP Address Binding Entries

In V800R015C00, The contents of IP Address Binding Entries are stored as user data management (UDM) data. Recovery of IP Address Binding Entries is a feature by using which the system restores the UDM data to the memory when the system is restarted. Users do not need to dial up again to bind IP addresses. Recovery of IP Address Binding Entries supports power-off recovery and non-power-off recovery. In non-power-off recovery, the memory space of the UDM data is not cleared. After the system is restarted, the system obtains the IP Address Binding Entries to implement IP address binding. Non-power-off recovery requires no configuration and the working principle is simpler than that of the power-off recovery. The following describes the power-off recovery scenario.

- When the system is running properly, the UDM data is periodically compressed and backed up on the server through FTP/TFTP/SFTP. During automatic system backup, DHCP, DHCPv6, SLAAC, and PPPoE dialup users are forbidden to go online or offline. This is intended to avoid data conflicts. If a device power failure occurs during automatic

system backup, the file stored on the server is incomplete. In such a case, the UDM data cannot be recovered after the system is restarted.

- When the system is restarted after a power failure, the system automatically downloads the backup data from the server and restores it after decompression. Because automatic download is performed during system startup, the upstream port may not be ready for automatic download and the download channel may not be available. In this case, automatic download cannot be smoothly carried out. The system makes attempts to download data from the server till the timeout time elapses. If no attempt is successful, the system does not make any further attempts. During automatic download, data recovery, and data download attempts, dialup users are not allowed to go online or offline. This is to avoid data conflicts. Once automatic data backup is disabled during data download or data download attempts, users can go online and offline. If automatic data backup is disabled during data recovery after data download, users can go online and offline only after the UDM data is recovered.
- When the system configured with active/standby servers is restarted due to a power failure, the system will try to download data from the active server first. If the active server is not available, the system will try the standby server. When the file downloaded from the active server fails to be verified or is not the latest, the system will not download data from the standby server.
- The lease time of the recovered UDM data may be different from that of the original UDM data when the system time is changed in the following conditions: before a device power failure occurs without any automatic data backup; after the system is restarted due to a power failure while the UDM data has not been completely recovered.
- If a device power failure occurs after you run the **active configuration system** command but before the first UDM data backup is complete, IP Address Binding Entries cannot be correctly recovered after the system is restarted.

27.6.3 Configuring Anti-IP Spoofing

This topic describes how to configure IP address binding and anti-IP spoofing to prevent malicious users from attacking the device or authorized users by forging the IP addresses of authorized users.

Context

IP address binding refers to binding an IP address to a service port. After the binding, the service port permits only the packet whose source IP address is the bound address to go upstream, and discards the packets that carry other source IP addresses.

Anti-IP spoofing is to dynamically trigger the IP address binding, preventing illegal users from stealing the IP address of legal users. When anti-IP spoofing is enabled, a user port is bound to an IP address after the user goes online. Then, the user cannot go online through this port by using other IP addresses, and any user cannot go online through other ports by using this IP address.

Procedure

- Configure the IP address binding.

  Run the **bind ip** command to bind an IP address to a service port.

  To permit only the users of certain IP addresses to access the system so that illegal users cannot access the system by using the IP addresses of legal users, configure the IP address binding.
- Configure anti-IP spoofing.

When the service flow binds to a VLAN service profile, anti-IP spoofing takes effect only when all its three levels are enabled.

When the service flow does not bind to any VLAN service profile, anti-IP spoofing takes effect only when two levels of anti-IP spoofing functions (the VLAN level function is not included) are enabled.

- Global function: Run the **security anti-ipspoofing** command to configure the global function. By default, the global function is disabled.
- The VLAN-level function and the service-port-level function are enabled by default. When the global function is enabled, anti-IP spoofing is effective to all the service flows of the system. To disable anti-IP spoofing for a service flow in this case, do as follows:
  - If the VLAN of the service flow is bound to a VLAN service profile and the VLAN service profile specifies that all of its service flows must disable anti-IP spoofing, disable the VLAN-level function for VLANs bound to this profile.
  - If only anti-IP spoofing of the service flow needs to be disabled, disable the service-port-level function.
- VLAN-level function:
  i. Run the **vlan service-profile** command to create a VLAN service profile and enter the VLAN service profile mode.
  ii. Run the **security anti-ipspoofing** command to configure the VLAN-level function. By default, the VLAN-level function is enabled.
  iii. Run the **commit** command to make the profile configuration take effect. The configuration of the VLAN service profile takes effect only after this command is executed.
  iv. Run the **quit** command to quit the VLAN service profile mode.
  v. Run the **vlan bind service-profile** command to bind the VLAN to the VLAN service profile configured in i.
- Service-port-level function: Run the **security anti-ipspoofing service-port** command to configure the service-port-level function. By default, the service-port-level function is enabled.

NOTE

When anti-IP spoofing is enabled after a user is already online, the IP address of this user is not bound by the system. As a result, the service of this user is interrupted, this user goes offline, and the user needs to go online again. Only the user who goes online after anti-IP spoofing is enabled can have the IP address bound.

- (Optional) Enable power-off recovery of IP address binding entries.

  If you want users to go on line without dialup after a device power failure occurs, configure this function.

  a. Run the **security user auto-backup enable** command to enable automatic data backup.
  b. Run the **file-server auto-backup udm** command to configure the auto-backup server.
  c. Run the **security user auto-backup period** command to configure the period for automatic data backup.
  d. Run the **security user auto-load timeout** command to configure the timeout parameters for automatic data download. The timeout parameters include the total timeout time and the interval between each download attempt. If download is not finished before the timeout time elapses, the system stops data download.

----End

Example

To bind IP address 10.1.1.245 to service port 2, that is, service port 2 permits only the packet whose source IP address is 10.1.1.245, do as follows:

```
huawei(config)#bind ip service-port 2 10.1.1.245
```

To enable anti-IP spoofing for service port 1 in service VLAN 10, do as follows:

```
huawei(config)#security anti-ipspoofing enable
huawei(config)#vlan service-profile profile-id 2
huawei(config-vlan-srvprof-2)#security anti-ipspoofing enable
  Info: Please use the commit command to make modifications take effect
huawei(config-vlan-srvprof-2)#commit
huawei(config-vlan-srvprof-2)#quit
huawei(config)#vlan bind service-profile 10 profile-id 2
huawei(config)#security anti-ipspoofing service-port 1 enable
```

27.7 IPv6 Anti-Spoofing

IPv6 anti-spoofing (in an IPv6 network topology) functions in a similar way to IPv4 anti-spoofing (in an IPv4 network topology). For details about specifications and principles, see 27.6 Anti-IP Spoofing. This topic describes the differences between IPv6 anti-spoofing and IPv4 anti-spoofing regarding their functions.

27.7.1 Principle

The procedure for binding IPv6 addresses dynamically is as follows:

1. After the dynamic IPv6 address learning function is disabled, the system monitors the users' (who may be using the DHCPv6 or SLAAC protocol) going online and offline processes. When a user goes online, the system dynamically obtains the user's source IPv6 address and binds the user's source IPv6 address to a service flow.
2. The system allows only the packets whose source IPv6 addresses are bound to the service flow to pass through.
3. When a user goes offline, the system unbinds the user's source IPv6 address from the service flow.

In an IPv6 network, users can obtain IPv6 addresses using stateless address autoconfiguration (SLAAC) or Dynamic Host Configuration Protocol version 6 (DHCPv6).

- In a network that uses SLAAC, the broadband network gateway (BNG) allocates IPv6 prefixes to users and the MA5600T/MA5603T/MA5608T dynamically binds these IPv6 prefixes to the service flow. To do so, the MA5600T/MA5603T/MA5608T obtains IPv6 prefixes allocated to the users from the router advertisement (RA) message sent by the BNG and dynamically generates IP address binding entries.
- In a network that uses DHCPv6, IP address binding is triggered by DHCPv6 packets when a user sends DHCPv6 packets to obtain an IP address. A DHCPv6 server may allocate one or more IPv6 addresses or IPv6 prefixes to the user through a DHCPv6 packet. The MA5600T/MA5603T/MA5608T obtains all the IPv6 addresses and prefixes allocated by the DHCPv6 server from the DHCPv6 packets received, and generates IP address binding entries.

In an IPv6 network, the MA5600T/MA5603T/MA5608T supports static binding of IPv6 addresses. The binding of an IPv6 address is different from that of an IPv4 address because of the differences between IPv6 and IPv4 address structures. In IPv6 binding, the MA5600T/MA5603T/MA5608T binds a variable-length IPv6 prefix to a service flow, whereas in IPv4 binding, the MA5600T/MA5603T/MA5608T binds a complete IPv4 address to a service flow.

27.8 User Account Anti-Forgery

Access devices use multiple protection methods to bind user accounts to access ports for authentication. This prevent user accounts against forgery and roaming on broadband networks.

User Account Anti-Forgery and Roaming Methods

Table 27-2 User account anti-forgery and roaming methods

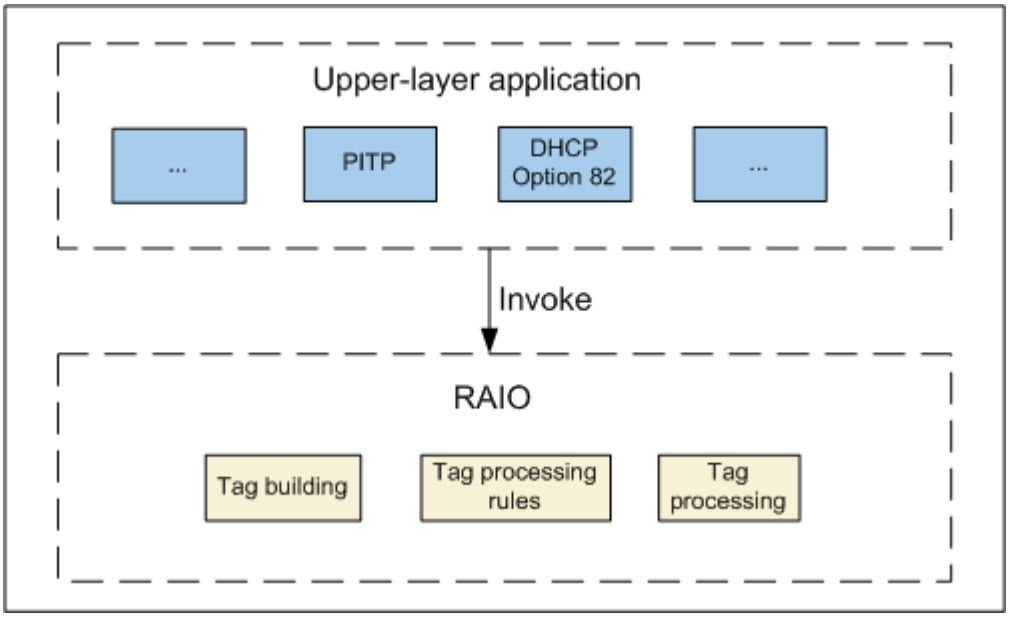
Security Feature	Description	Application Scenario
27.8.1 RAIO	Relay agent info option (RAIO) encapsulates the obtained user access information into protocol packets in various formats to provide the users' physical locations to the broadband remote access server (BRAS) or Dynamic Host Configuration Protocol (DHCP) server. This implements secure access and management for user accounts.	RAIO applies in DHCP Option 82 and PITP features. For the relationship between the RAIO feature and the DHCP option 82 and PITP features, see Figure 27-7. The RAIO configuration varies according to the feature to which RAIO applies. For instructions about how to configure RAIO, see the configuration procedures of the preceding features.
27.8.2 DHCP Option 82	DHCP Option 82 is a user security mechanism, which encapsulates the user access information obtained by access devices through RAIO into the Option 82 field of the DHCP request packets sent from a user. The data is encapsulated in the format specified by customers. This facilitates the upper-layer authentication server to authenticate users and prevents user account theft and roaming.	The DHCP relay agent inserts the DHCP Option 82 information into the DHCP packets sent from a DHCP client to identify the access location of the DHCP client. Therefore, the DHCP Option 82 function applies on DHCP relay networks.
27.8.3 PITP	The PITP protocol provides the information about	PITP is available in PITP P mode and PITP V mode and used in PPPoE.

Security Feature	Description	Application Scenario
	physical ports of access users to the BRAS. After obtaining the physical port information, the BRAS authenticates the binding between the user accounts and the physical ports, thereby preventing user account theft and roaming.	• The PITP V mode uses the virtual broadband access server (VBAS) protocol. • The PITP P mode is also named PPPoE+ mode, which is widely used.

Relationship Between the RAIO Feature and Other Features

Figure 27-7 shows the interaction between the RAIO feature and upper-layer features. The upper-layer application modules invoke the RAIO feature. The RAIO feature adapts to each upper-layer application module.

Figure 27-7 Relationship between the RAIO feature and other features



27.8.1 RAIO

Relay agent info option (RAIO) implements secure access and management for user accounts.

Introduction

Feature Value

Access devices obtain the user access information and send it to the core network devices. Based on user access information, the core network devices implement authentication, accounting, secure control, and bandwidth limitation before providing data services for a user.

In network applications, multiple user secure protocols are used to define user access information, such as the Policy Information Transfer Protocol (PITP), Dynamic Host Configuration Protocol (DHCP) Option 82, DHCPv6 Option, and Access Node Control Protocol (ANCP). The user access information needs to be processed by all access devices using a public feature. The relay agent info option (RAIO) feature has been developed to meet this requirement.

Function

RAIO encapsulates the obtained user access information into protocol packets in various formats to provide the users' physical locations to the broadband remote access server (BRAS) or Dynamic Host Configuration Protocol (DHCP) server. This implements secure access and management for user accounts.

Standard and Protocol Compliance

The RAIO feature complies with *TR101 Migration to Ethernet-Based DSL Aggregation*.

Basic Concepts

Tag Types

Tag type is specified in type-length-value (TLV). Relay agent info option (RAIO) supports the following tag types:

- Circuit ID (CID): used to identify a subscriber line. A CID contains the subrack ID, slot ID, port ID, and VPI/VCI (VPI is the abbreviated form of virtual path identifier, and VCI is the abbreviated form of virtual channel identifier). Carriers use a CID to identify a user and control the user access. A CID is in the TLV format and has a fixed value of 1.
- Remote ID (RID): used to identify a remote user. An RID contains the remote flow description, phone number, and user name. An RID is in the TLV format and has a fixed value of 2.
- Option 18: an option defined by Dynamic Host Configuration Protocol version 6 (DHCPv6) and filled with a CID.
- Option 37: an option defined by DHCPv6 and filled with an enterprise ID and an RID.
- Information (INFO): dedicated for working objects in Policy Information Transfer Protocol (PITP) V mode and filled with a CID.
- Sub-option: refers to sub-option 0x81 to sub-option 0x91, which are described in Table 27-3.

NOTE
Only the PITP V mode and DHCP Option 82 support sub-options.

Table 27-3 Description of sub-options

Sub-option	Description
Sub-option 81	Actual upstream rate in the activated state.
Sub-option 82	Actual downstream rate in the activated state.
Sub-option 83	Minimum upstream data rate.
Sub-option 84	Minimum downstream data rate.
Sub-option 85	Minimum upstream reserved rate.
Sub-option 86	Minimum downstream reserved rate.
Sub-option 87	Maximum upstream data rate.
Sub-option 88	Maximum downstream data rate.
Sub-option 89	Minimum upstream rate in low power state.
Sub-option 8A	Minimum downstream rate in low power state.
Sub-option 8B	Maximum upstream interleave delay.
Sub-option 8C	Actual upstream interleave delay.
Sub-option 8D	Maximum downstream interleave delay.
Sub-option 8E	Actual downstream interleave delay.
Sub-option 8F	Line status.
Sub-option 90	Subscriber line type and data encapsulation type.
Sub-option 91	Line transmission type.

RAIO Modes

Each RAIO mode defines various tag formats. A RAIO mode can be pre-defined or user-defined. A pre-defined mode focuses on customers' requirements, and a user-defined mode features flexibility. Each RAIO mode can define various tag formats. In pre-defined mode, a tag format is pre-defined.

Table 27-4 RAIO modes

Mode		Description
Pre-defined mode	Standard pre-defined mode	Is defined by standard organizations. dslforum-default and broadband forum (BBF) standard modes are supported. dslforum-default is the default mode defined by the DSL forum. BBF is defined by the Broadband Forum and complies with the TR156 standard.
	Customer pre-defined mode	Is customized based on carriers' requirements. In this mode, tag formats are defined by carriers. For example, cntel-xpon, cntel, ft, and ti are customer pre-defined modes.

Mode		Description
	Device pre-defined mode	Is a universal mode defined by devices. For example, common is a pre-defined mode.
User-defined mode		Features flexible configurations. The access type, keyword, separator, length, and option sign are configurable.

NOTE
The tag formats can be configured by running the **raio-mode** command.

Tag Formats in Pre-defined Mode

A CID format identifies the global attributes of a device. An RID format identifies the access information (not global information) of a user. The CID and RID formats vary according to the RAIO mode and access mode. Table 27-5 lists the CID and RID formats in various RAIO modes and access modes.

NOTE
The following table lists only the tag formats in standard pre-defined and device pre-defined modes. The customer pre-defined mode is based on customers' requirements. In this mode, the tag formats are customized and therefore are not described in the following table.

Table 27-5 Tag formats in various access modes

RAIO Mode	Access Mode	CID Format	RID Format
dslforum-default (standard pre-defined mode)	ATM	anid atm slot/port: vpi.vci	None
	VDSL or LAN	• Multi-service based on VLANs: anid eth slot/port: flowpara • Others: anid eth slot/port: vlanid	None
	xPON	• Multi-service based on VLANs: anid xpon frame/slot/0/port: gemport.ontid.flowpara • Others: anid xpon frame/slot/0/port: gemport.ontid.vlanid	None
	EoC	anid eth slot/port: vlanid	None

BBF (standard pre-defined mode)	ATM	• In digital subscriber line access multiplexer (DSLAM) scenarios: anid atm slot/port: vpi.vci • In fiber to the x (FTTx) scenarios: anid atm slot/port/onuid/slot/port: vpi.vci	• In digital subscriber line access multiplexer (DSLAM) scenarios: empty • In fiber to the x (FTTx) scenarios: ONT label
	VDSL or LAN	• In DSLAM scenarios: anid eth slot/port: [vlan-id] • In FTTx scenarios: anid eth slot/port/onuid/slot/port: [vlan-id]	• In DSLAM scenarios: empty • In FTTx scenarios: ONT label
Common (device pre-defined mode)	ATM	anid atm frame/slot/subslot/port: vpi.vci	None
	VDSL or LAN	anid eth frame/slot/subslot/port: Emptyvlanid	None
	xPON	anid xpon frame/slot/subslot/port: ontid.gemport.vlanid	None
	EoC	anid eoc frame/slot/subslot/port: cnuid	None
	DOCSIS	anid docsis frame/slot/subslot/port	splabel
xDSL port rate (device pre-defined mode) **NOTE** In this mode, a CID is formed by adding the upstream and downstream rates of an ADSL port in the activated state to the	ATM	anid atm frame/slot/subslot/port: vpi.vci%up: uprate down: dnrate **NOTE** • **%**: followed by the rates in the activated state. • **up**: indicates the upstream rate in	User-defined

end of the default CID format. Only ADSL2+ boards support the xDSL port rate mode.		the activated state. • **down**: indicates the downstream rate in the activated state.	
	VDSL or LAN	anid eth frame/slot/subslot/port: vlanid%up: uprate down: dnrate	User-defined
	xPON	anid xpon frame/slot/subslot/port: ontid.gemport.vlanid%up: uprate down: dnrate	User-defined
	EoC	anid eth frame/slot/subslot/port: vlanid%up: uprate down: dnrate	User-defined
port-userlabel (device pre-defined mode) **NOTE** In this mode, a CID carries the label of a user port (user-defined port description) with a maximum length of 32 bytes in addition to the information required by the default format. An RID carries the label of a user port.	ATM	anid atm slot/port: vpi.vci	plabel
	VDSL or LAN	anid eth slot/port: vlanid	plabel
	xPON	anid xpon frame/slot/subslot/port: ontid.gemport.vlanid	plabel
	EoC	anid eth slot/port: vlanid	plabel
service-port-userlabel (device pre-defined mode) **NOTE** In service-port-userlabel mode, an RID carries user-defined service flow description, which can be configured by running the **service-port desc** command. In port-userlabel mode, an RID carries the label of a user port.	ATM	anid atm slot/port: vpi.vci	splabel
	VDSL or LAN	• Multi-service based on VLANs: anid eth slot/port: flowpara • Others: anid eth slot/port: vlanid	splabel
	xPON	• Multi-service based on VLANs: anid xpon frame/slot/0/port: gemport.ontid.flo	splabel

		wpara • Others: anid xpon frame/slot/0/port: gemport.ontid.vlanid	
	EoC	anid eth slot/port: vlanid	splabel

In the preceding table:

- **anid** is a character string that identifies an access node. It can contain any characters but a space or separator is not recommended. The BBF mode does not allow a space in **anid**. Fill in **anid** by following the rules:
  - If **anid** has been configured, use the configured value.
  - If **anid** has not been configured but the device name has been configured, use the device name.
  - If neither **anid** nor the device name has been configured, use the MAC address of the device.
  - In BBF mode, if user packets carry a use-side VLAN, **vlan-id** is the ID of the VLAN.

Tag Formats in User-defined Mode

The CID and RID formats are customized in user-defined mode. The following describes the syntax rules for the user-defined mode.

- Only the keyword and separator sets defined in the Router can be parsed. The keyword set contains the minimum keyword set defined by TR101 and the keyword set extended by the Router. For details, see Table 27-6.
- Maximum width

  The maximum width refers to the maximum number of columns for a keyword. The maximum widths of keywords specified in the Router are greater than the maximum width defined in TR101. The reason is that the actual maximum width required by some manufacturers is greater than the maximum width defined in TR101. The maximum width of **anid** is determined by the maximum character string length (50 characters) supported by the Router.
- Configurable width

  The number of columns for a keyword can be configured. The Router automatically adds 0s to the beginning of the number of used columns if the number of columns used by a keyword is less than the configured width. The syntax is "keyword+0+m", where "m" indicates the number of columns used by a keyword. For example, "slot03" indicates that the number of columns used by the slot keyword is 3. Therefore, if a slot occupying two columns, it is displayed as 002 in a packet. "m" must be less than or equal to the maximum width. If the actual number of columns is greater than "m", "m" is displayed.

Table 27-6 User-defined keyword set

Keyword	Description	Whether the Width Is Configurable	Maximum Width
anid	Name of an access node	No	63
anip	IP address of an access node	No	15
eth	Ethernet access mode	No	3
accesstype	User access type, which takes effect only on xPON lines	No	4
atm	ATM access mode	No	3
xpon	xPON access mode	No	4
chassis	Cabinet ID of an access node	Yes	4
rack	Rack ID of an access node	Yes	4
frame	Subrack ID	Yes	4
slot	Slot ID	Yes	4
logicalslot	Logical slot ID Only service boards have logical slot IDs. An idle service board slot also has a logical slot ID. Non-service board slots, such as the slots for control boards, upstream boards, and power boards, do not have logical slot IDs. Logical slot IDs in CIDs and RIDs are continuous service board slot IDs starting from 1.	Yes	4
subslot	Daughter board ID, which is filled with 0	Yes	4
port	Port ID	Yes	4
port+1	Port ID plus 1. The port ID in a CID and RID is the actual port ID plus 1. If a working object works in Access Node Control Protocol (ANCP) mode, both the **ancp port begin** command and the **port+1** keyword can take effect. That is, if the start port ID is set to 1 by running the **ancp port begin** command, the port ID in the CID and RID is the actual port ID plus 1.	Yes	4
cvlanid	User-side VLAN ID If services carried over service ports are identified by user-side VLAN IDs, the value of this keyword is the user-side VLAN ID on a service port. If services carried over service ports are not identified by user-side VLAN IDs, the value of this keyword is null.	Yes	4

Keyword	Description	Whether the Width Is Configurable	Maximum Width
vlanid	VLAN ID If services carried over service ports are identified by user-side VLAN IDs, the value of this keyword is the user-side VLAN ID on a service port. If services carried over service ports are not identified by user-side VLAN IDs, the value of this keyword is the ID of the network-side VLAN ID.	Yes	4
priority	Priority of the traffic profile for service ports when Layer-2 PPPoE and Dynamic Host Configuration Protocol (DHCP) Option 82 are enabled	Yes	4
plabel	Description of a user port	No	32
splabel	Description of a service port The description can be configured by running the **service-port desc** command.	No	64
sprlabel	Description of the remote port connected to a service port The description can be configured by running the **service-port remote-desc** command.	No	64
bslot	Broadband remote access server (BRAS) slot ID	Yes	4
bsubslot	BRAS sub-slot ID	Yes	4
bport	BRAS port ID	Yes	4
bporttype	BRAS access mode	Yes	4
8021p	VLAN priority	Yes	4
xpi	• If the attribute of a network-side VLAN is stacking, the value of this keyword is the ID of the network-side VLAN. • If the attribute of a network-side VLAN is not stacking, this keyword has a fixed value of 4096.	Yes	4
xci	• If the attribute of a network-side VLAN is stacking, the value of this keyword is the label of the service port. • If the attribute of a network-side VLAN is not stacking, the value of this keyword is the ID of the network-side VLAN.	Yes	5
axpi (Used in ATM	VPI	Yes	4

Keyword	Description	Whether the Width Is Configurable	Maximum Width
access mode)			
axpi (Used in Ethernet and xPON access modes)	Network-side VLAN ID	Yes	4
axci (Used in ATM access mode)	VCI	Yes	5
axci (Used in Ethernet and xPON access modes)	When the attribute of the network-side VLAN is stacking: • If services carried over service ports are identified by user-side VLAN IDs, the value of this keyword is the user-side VLAN ID on a service port. • If services carried over service ports are not identified by user-side VLAN IDs, the value of this keyword is the label of a service port. When the attribute of the network-side VLAN is not stacking: • If services carried over service ports are identified by user-side VLAN IDs, the value of this keyword is the user-side VLAN ID on a service port. • If services carried over service ports are not identified by user-side VLAN IDs, the keyword has a fixed value of 4096.	Yes	5
gem-index	GPON encapsulation mode (GEM) index of an xPON line	Yes	4
gemport	GEM port ID of an xPON line	Yes	4
uprate	Upstream rate of an ATM or Ethernet port (The Ethernet port is only a PTM port.)	Yes	10
dnrate	Downstream rate of an ATM or Ethernet port (The Ethernet port is only a PTM port.)	Yes	10
ontid	Optical network unit (ONU) ID of an xPON line	Yes	4

Keyword	Description	Whether the Width Is Configurable	Maximum Width
ontid+1	Optical network terminal (ONT) ID plus 1 The ONT ID in a CID and RID is the actual ONT ID plus 1.	Yes	4
ontlabel	ONT label, which takes effect only on xPON lines	No	63
ontporttype	ONT port type, which can be: • **1**: indicates Ethernet ports, including those of the wireless local area network (WLAN) type. • **2**: indicates POTS ports.	No	63
ontportid	ONT port number **NOTE** The value of this keyword is 0 for Ethernet ports of the WLAN type and for POTS ports on a single voice access gateway (VAG).	No	63
onuid	ONU ID of an xPON line The value of this keyword is the MAC address of an EPON ONU or the serial number (SN) of a GPON ONU. The keyword is padded with 0s at the beginning if it is shorter than 24 bits.	Yes	24
onutag	Data after "ANID accesstype" in the tag carried by a message transmitted or received by an ONU, which takes effect only on xPON lines For example, in "ANID accesstype slot/port: vlanid", the value of this keyword is "lot/port: vlanid".	No	255
SN	ONT SN	No	16
0002	Fixed filling	Yes	4
up	Fixed filling	Yes	2
down	Fixed filling	Yes	4
vpi	VPI of an ATM line	Yes	4
vci	VCI of an ATM line	Yes	4
ge	Fixed filling	Yes	2
Plaintext	Identified using a pair of quotation marks ("") in RAIO mode The plaintext consists of letters, digits, spaces, and the following special characters: + * - / | . : < > [ ] , # @ $ % !	Yes	N/A
Option	Displayed in the format of square brackets ([]) to identify	N/A	N/A

Keyword	Description	Whether the Width Is Configurable	Maximum Width
sign	optional keywords in RAIO mode Only **cvlanid** is optional.		

- The CID format character string must contain **anid**.
- The port type keyword identifies the format of a port type.
- A format character string cannot contain keywords that are used for different port types. For example, **vpi** and **gemport**, or **eth** and **vci** are invalid in a character string.
- If a port type is not specified, the CID and RID are empty.
- A separator identifies a character string in RAIO mode and will be added to a CID and RID. A separator can be a space, full stop (.), colon (:), slash (/), hyphen (-), percent (%), comma (,), semicolon (;), number sign (#), or exclamation point (!).
- The length of a tag character string contains 1-127 lowercase characters.
- A CID character string must contain **anid**.
- **anid** must be in front of the port type keyword.
- The following separators are used for parsing **anid** in downstream packets: all separators in front of **anid** in a CID character string, RAIO separators (if available) in **anid**, and the first separator following **anid**.

The following provides an example of a tag format in user-defined mode.

The following configurations are used as an example:

- Device name: DSLAM01
- Slot ID: 3
- Port ID: 15
- VPI: 0
- VCI: 35
- Priority: 6

The user-defined CID is "anid atm slot/port: vpi.vci%priority". Therefore, the generated character string is "dslam01 atm 3/15: 0.35%6".

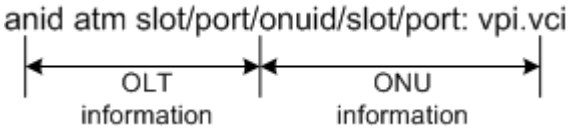
Rebuilding

Enable rebuilding if a tagged packet must contain ONU and optical line terminal (OLT) access information in FTTx scenarios. On the OLT, the ONU and OLT access information must be integrated and rebuilt in the format defined by the RAIO mode. Ensure that the RAIO modes configured on the ONU and the OLT are the same, and rebuilding is enabled on the OLT.

Only the working objects of the PITP P mode and DHCP Option 82 support rebuilding. Run the **pitp** and **dhcp option82** commands to enable rebuilding, respectively.

The rebuilding function must be used in BBF, vnpt, or user-defined mode. The reason is that in the three modes, packets contain two tags for carrying ONU and OLT access information, as shown in Figure 27-8.

Figure 27-8 Two tags



The differences in rebuilding the tag format in the DSLAM network and the FTTx network are as follows:

- In DSLAM scenarios, rebuilding is disabled generally. If rebuilding is enabled, the device selects the ATM or Ethernet type to rebuild tagged packets according to the user access type. The tagged packets contain only the OLT access information.

  Consider the BBF mode as an example. In DSLAM scenarios, the tag format is "anid atm slot/port: vpi.vci". The tagged packets contain only the OLT access information.
- In FTTx scenarios, if rebuilding is enabled, the device rebuilds tagged packets according to the xPON type. The tagged packets contain the ONU and OLT access information.
- If rebuilding is disabled, the device rebuilds tagged packets according to the Ethernet type. The tagged packets contain only the OLT access information.

27.8.2 DHCP Option 82

Dynamic Host Configuration Protocol (DHCP) Option 82 is a user security mechanism. In this mechanism, a user's physical location information is added to the Option 82 field of the request packets sent by the user. This facilitates the upper-layer authentication server to authenticate users.

Introduction

Feature Value

The widely used Dynamic Host Configuration Protocol (DHCP) does not support authentication or security mechanisms. Therefore, DHCP encounters many security issues in network applications compared with Peer-to-Peer Protocol (PPP), such as frequent DHCP broadcast, DHCP IP address exhaustion and attacks, IP address spoofing, MAC address spoofing, and user ID spoofing. In addition, DHCP clients cannot be managed in a unified manner. To resolve these issues, RFC3046 defines the "DHCP Relay Agent Information Option" field in DHCP packets. The ID of the field is 82. Therefore, the field is named DHCP Option 82. A DHCP client sends DHCP packets to the DHCP server to request for an IP address. If the DHCP packets carry the Option 82 field, the DHCP server verifies the DHCP client according to the Option 80 field. This ensures the user access security.

Function

DHCP Option 82 is a user security mechanism, which encapsulates the user access information obtained by access devices through relay agent info option (RAIO) into the Option 82 field of the DHCP request packets sent from a user. The data is encapsulated in the

format specified by customers. This facilitates the upper-layer authentication server to authenticate users and prevents user account theft and roaming.

NOTE
For details about the formats of DHCP Option 82 packets, see 27.8.1 RAIO.

Standard and Protocol Compliance

The DHCP Option 82 feature complies with:

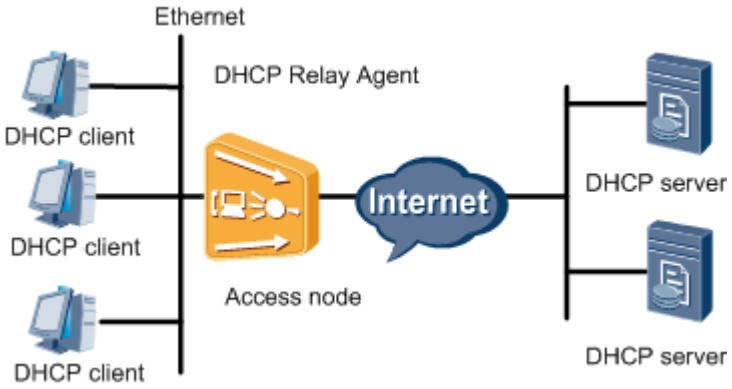
- *RFC2131 Dynamic Host Configuration Protocol*
- *RFC3046 DHCP Relay Agent Information Option*

Network Application

The DHCP relay agent inserts the Dynamic Host Configuration Protocol (DHCP) Option 82 information into the DHCP packets sent from a DHCP client to identify the access location of the DHCP client. Therefore, the DHCP Option 82 function applies on DHCP relay networks.

Figure 27-9 shows the typical DHCP relay networking.

Figure 27-9 Typical DHCP relay networking



The preceding figure involves the following roles:

- DHCP client: a device that dynamically obtains an IP address or other network configuration parameters.
- DHCP relay agent: a relay agent that adds the Option 82 information to the request packets sent from a DHCP client to the DHCP server and forwards the request packets to the DHCP server to obtain an IP address and other network configuration parameters if the DHCP client and the DHCP server connect to different links. This prevents the deployment of a DHCP server for each link, thereby reducing deployment costs and facilitating centralized management.
- DHCP server: a device that assigns IP addresses and other network configuration parameters to DHCP clients.

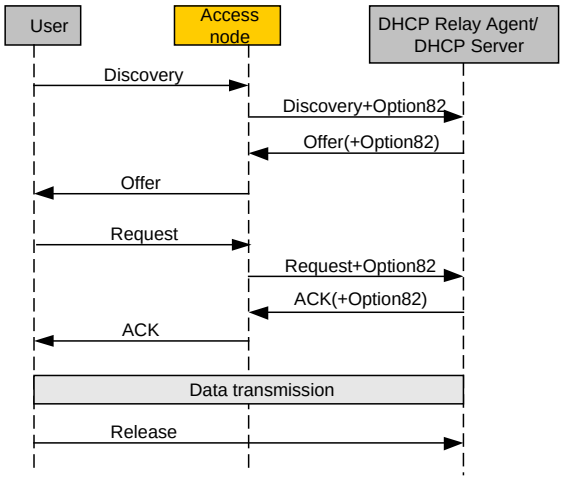
Principles

Basic Principles

Dynamic Host Configuration Protocol (DHCP) Option 82 can be enabled or disabled at global, port, service port, or VLAN level. It takes effect only after being enabled at the four levels.

Figure 27-10 shows the DHCP interactive process after DHCP Option 82 is enabled.

Figure 27-10 DHCP interactive process after DHCP Option 82 is enabled



If the request packets sent from a DHCP client to the DHCP server pass through the DHCP relay agent, the DHCP relay agent adds Option 82 data to the request packets. The DHCP Option 82 function enables the DHCP server to obtain the IP addresses of the DHCP client and relay agent. By working with software, the DHCP Option 82 function implements accounting and limited IP address assignment.

1. The DHCPv4 client broadcasts request packets during initialization.
2. If no DHCP server is available, the DHCP relay agent checks whether the request packets contain the Option 82 field and performs follow-up operations.

NOTE

If a DHCP server is available, the DHCP client obtains an IP address from the server.

   - If the Option 82 field is available in the request packets, the DHCP relay agent replaces the Option 82 field with that of itself or retains the Option 82 field according to the configured policy listed in Table 27-7. Then, the DHCP relay agent sends the request packets to the DHCP server.
   - If no Option 82 field is available in the request packets, the DHCP relay agent adds the Option 82 field to the packets and sends the packets to the DHCP server. In this case, the request packets contain the MAC address of the switch port connected to the DHCP client, ID of the VLAN to which the switch port belongs, and MAC address of the DHCP relay agent.

3. After receiving the DHCP request packets sent from the DHCP relay agent, the DHCP server records the information contained in the Option 82 field and sends the packets carrying DHCP configuration and Option 82 data to the DHCP relay agent.

4. After receiving the packets sent from the DHCP server, the DHCP relay agent processes the Option 82 data in the packets according to the policy shown in Figure 27-11 and sends the processed packets to the DHCPv4 client.

Policies used by access devices to process DHCP packets

- By default, the global DHCP option82 is disabled while DHCP option82 is enabled for a port. If DHCP option82 is disabled globally, even if DHCP option82 is enabled on a port, no vendor tag is added to the DHCP packets sent from the port. Only when DHCP option82 is enabled globally and on a port, vendor tags are added to the DHCP packets sent from the port.
- Run the **dhcp-option82 permit-forwarding service-port** command to configure whether a service port allows user-side DHCP packets to carry the Option 82 information. Table 27-7 lists the policies used by access devices to process user-side DHCP packets.

Figure 27-11 Network-side DHCP packet processing policy used by access devices

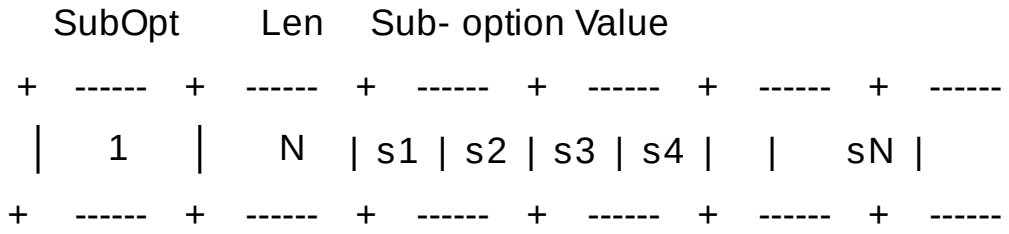


Table 27-7 Policies used by access devices to process user-side DHCP packets

Status of DHCP Option 82 in Global and VLAN Service Profile Mode	Port and Service Port Allowing User-Side DHCP Packets to Carry the Option 82 Data	Whether DHCP Packets Carry the Option 82 Data	User-Side DHCP Packet Processing Policy Used by the Access Device
enable	**enable**	Yes	The access device removes the Option 82 data carried in the user-side DHCP packets and adds the local user information to the packets.
		No	The access device adds the local user information to the user-side DHCP packets.
	disable	Yes	The access device discards the user-side DHCP packets.
		No	The access device adds the local user information to the user-side DHCP packets.

Status of DHCP Option 82 in Global and VLAN Service Profile Mode	Port and Service Port Allowing User-Side DHCP Packets to Carry the Option 82 Data	Whether DHCP Packets Carry the Option 82 Data	User-Side DHCP Packet Processing Policy Used by the Access Device
forward	**enable**	Yes	The access device forwards the network-side DHCP packets.
		No	The access device adds the local user information to the user-side DHCP packets.
	disable	Yes	The access device discards the user-side DHCP packets.
		No	The access device adds the local user information to the user-side DHCP packets.
rebuild	**enable**	Yes	The access device re-tags the user-side DHCP packets.
		No	The access device re-tags the user-side DHCP packets.
	disable	Yes	The access device discards the user-side DHCP packets.
		No	The access device re-tags the user-side DHCP packets.
disable	**enable** or **disable**	Yes or No	The access device forwards the network-side DHCP packets.

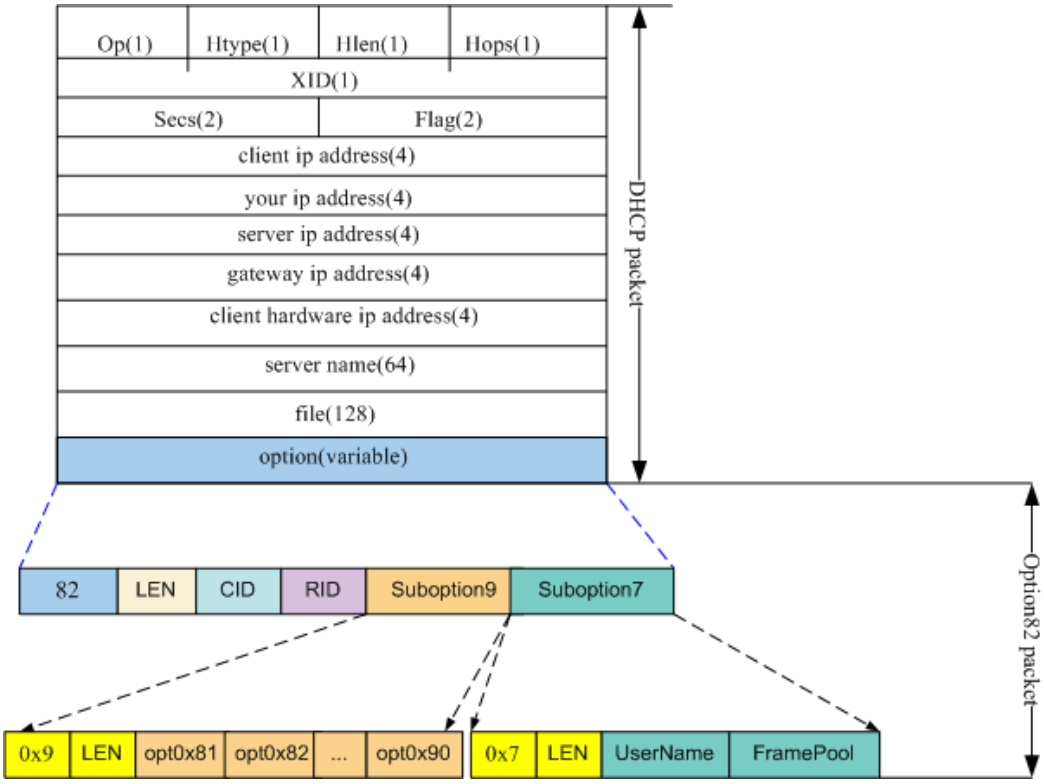
Format of a DHCP Option 82 Packet

Configure the format of a DHCP Option 82 packet before configuring the DHCP Option 82 function. The following section describes how to configure the packet format.

Each DHCP packet contains a variable option field with ID 82. Therefore, this field is named Option 82. Option 82 is extended for DHCP option, DHCP relay agent information option (RAIO).

Figure 27-12 shows the format of a DHCP Option 82 packet.

Figure 27-12 Format of a DHCP Option 82 packet



In RAIO (see 27.8.1 RAIO), the Option 82 field contains not only circuit ID (CID) and remote ID (RID) but also sub-option 7 and sub-option 9. Sub-option 9 contains other sub-options.

For details about sub-options, see 27.8.1 RAIO.

Configuring Anti-Theft and Roaming of DHCPv4 User Accounts Using DHCP Option 82

Context

When configuring the DHCP Option 82 function, you must configure relay agent info option (RAIO). RAIO can be configured in global or profile mode. The RAIO configuration in profile mode takes precedence over that in global mode.

Table 27-8 lists the default settings of DHCP Option 82.

Table 27-8 Default settings of DHCP Option 82

Parameter	Default Setting
Dynamic Host Configuration Protocol (DHCP) Option 82 function	Global status: disabled Port-level status: enabled VLAN-level status: enabled Service port-level status: enabled

Parameter	Default Setting
DHCP sub-option 7 function	Disabled
DHCP sub-option 90 function	Disabled
Whether a service port allows user-side DHCP packets to carry the Option 82 information	No

Procedure

Configure RAIO in global mode.

- Run the **raio-mode** ***mode*** **dhcp-option82** command to configure the RAIO mode.
- (Optional) If the RAIO mode is user-defined, run the **raio-format dhcp-option82** command to configure the RAIO format.
  - In user-defined RAIO mode, configure circuit ID (CID) and remote ID (RID).
  - In non-user-defined RAIO modes, the RAIO format is fixed and does not require configuration.
  - If an access mode is specified, the configured RAIO format takes effect only in this access mode. If no access mode is specified, the configured RAIO format takes effect in all access modes.

NOTE
For details about the RAIO format, see the **raio-format** command.

Step 1 Configure RAIO in profile mode.

1. Run the **raio-profile** command to create a RAIO profile.
2. Run the **raio-mode** ***mode*** **dhcp-option82** command to configure the RAIO mode.
3. (Optional) If the RAIO mode is user-defined, run the **raio-format dhcp-option82** command to configure the RAIO format.
   - In user-defined RAIO mode, configure CID and RID.
   - In non-user-defined RAIO modes, the RAIO format is fixed and does not require manual configuration.
   - If an access mode is specified, the configured RAIO format takes effect only in this access mode. If no access mode is specified, the configured RAIO format takes effect in all access modes.

NOTE
For details about the RAIO format, see the **raio-format** command.

4. Run the **quit** command to quit the RAIO profile mode.
5. Run the **vlan bind raio-profile** command to bind the RAIO profile created in Step 2.1 to a VLAN.

Step 2 (Optional) Run the **dhcp-option82 permit-forwarding service-port** command to configure a service port to allow user-side DHCP packets to carry the Option 82 information.

The DHCP relay agent adds the device name, subrack ID, slot ID, and port ID to the Option 82 field of DHCP packets to generate tagged packets. Then, the MA5600T/MA5603T/MA5608T forwards the tagged packets. Otherwise, the MA5600T/MA5603T/MA5608T discards the tagged packets.

Step 3 Enable the DHCP Option 82 function.

The DHCP Option 82 function can be enabled or disabled at four levels: global, port, VLAN, and service port. It takes effect only after being enabled at the four levels.

1. Run the **dhcp option82** command to enable the DHCP Option 82 function globally.
2. Run the **dhcp option82 board** or **dhcp option82 port** command to enable the DHCP Option 82 function at port level.
3. Enable the DHCP Option 82 function at VLAN level.
   a. Run the **vlan service-profile** command to create a VLAN service profile.
   b. Run the **dhcp option82** command to enable the DHCP Option 82 function at VLAN level.
   c. Run the **commit** command to make the profile configuration take effect.

      The configuration of the VLAN service profile takes effect only after this command is executed.
   d. Run the **quit** command to quit the VLAN service profile mode.
   e. Run the **vlan bind service-profile** command to bind the VLAN service profile created in Step 4.3.a to a VLAN.
4. Run the **dhcp option82 service-port** command to enable the DHCP Option 82 function at service port level.

Step 4 (Optional) Enable the sub-option function.

1. Run the **dhcp sub-option7** command to enable the sub-option 7 function.
2. Run the **dhcp sub-option90** command to enable the sub-option 90 function.
3. Run the **raio sub-option** command to configure the sub-option 0x81 to sub-option ox91 functions of sub-ption 90 to support the reporting of TR101 line parameters.

----End

Result

After the configuration, you can obtain the IP address using DHCP and connect to the Internet.

Example

The following configurations are used as an example to configure DHCP Option 82 to enhance user security:

- RAIO configuration mode: global
- RAIO mode: user-defined
- Ethernet access mode:
  - CID format: eth
  - Subrack ID/slot ID/sub-slot ID/port ID: vlanid
- xPON access mode:
  - CID format: xpon
  - Subrack ID/slot ID/sub-slot ID/port ID: ontid.vlanid
- RID format: labels of service ports

```
huawei(config)#raio-mode user-defined dhcp-option82
huawei(config)#raio-format dhcp-option82 cid eth anid eth
frame/slot/subslot/port:vlanid
huawei(config)#raio-format dhcp-option82 cid xpon anid xpon
frame/slot/subslot/port:ontid.vlanid
huawei(config)#raio-format dhcp-option82 rid eth splabel
huawei(config)#raio-format dhcp-option82 rid xpon splabel
huawei(config)#dhcp option82 enable
```

The following configurations are used as an example to configure DHCP Option 82 in VLAN 11 to enhance user security:

- RAIO profile index: 10
- RAIO mode: port-userlabel

```
huawei(config)#raio-profile index 10
huawei(config-raio-profile-10)#raio-mode port-userlabel dhcp-option82
huawei(config-raio-profile-10)#quit
huawei(config)#vlan bind raio-profile 11 index 10
huawei(config)#dhcp option82 enable
```

27.8.3 PITP

This topic describes the Policy Information Transfer Protocol (PITP) and PITP working principles. PITP is available in the PITP P mode and PITP V mode.

Introduction

Feature Value

The widely used PPPoE protocol is prone to user account theft and roaming due to the lack of methods for identifying and binding users' physical locations. The user account theft and roaming are hot complaint issues for broadband networks. The Policy Information Transfer Protocol (PITP) supported by Huawei MA5600T/MA5603T/MA5608T can resolve these issues. The PITP protocol provides the information about physical ports of access users to the broadband remote access server (BRAS). After obtaining the physical port information, the BRAS authenticates the binding between the user accounts and the physical ports, thereby preventing user account theft and roaming.

Function

PITP transfers users' physical port information from access devices to the BRAS in Layer 2 point-to-point (P2P) mode. PITP is available in the PITP P mode and PITP V mode.

- In PITP P mode, during the PPPoE negotiation between an access user and the BRAS, the MA5600T/MA5603T/MA5608T uses the RAIO function to add the physical port information of the access user to a PPPoE packet and sends the PPPoE packet to the BRAS.

  NOTE
  The PITP P mode is also named PPPoE+ mode, which is widely used.

- In PITP V mode, during the PPPoE negotiation between an access user and the BRAS, the BRAS proactively sends a VBRAS request packet to the MA5600T/MA5603T/MA5608T, asking the MA5600T/MA5603T/MA5608T to report the information about the physical port of the access user. The

MA5600T/MA5603T/MA5608T responds to the VBRAS request packet. Then, the MA5600T/MA5603T/MA5608T uses the relay agent info option (RAIO) function to add the obtained physical port information to a PPPoE packet in the format specified by the customer and sends the PPPoE packet to the BRAS.

NOTE

The PITP V mode uses the virtual broadband access server (VBAS) protocol, which is a non-standard protocol proposed by China Telecom. Therefore, the PITP V mode is used by China Telecom.

Standard and Protocol Compliance

The PITP feature complies with *RFC2516 PPP over Ethernet*.

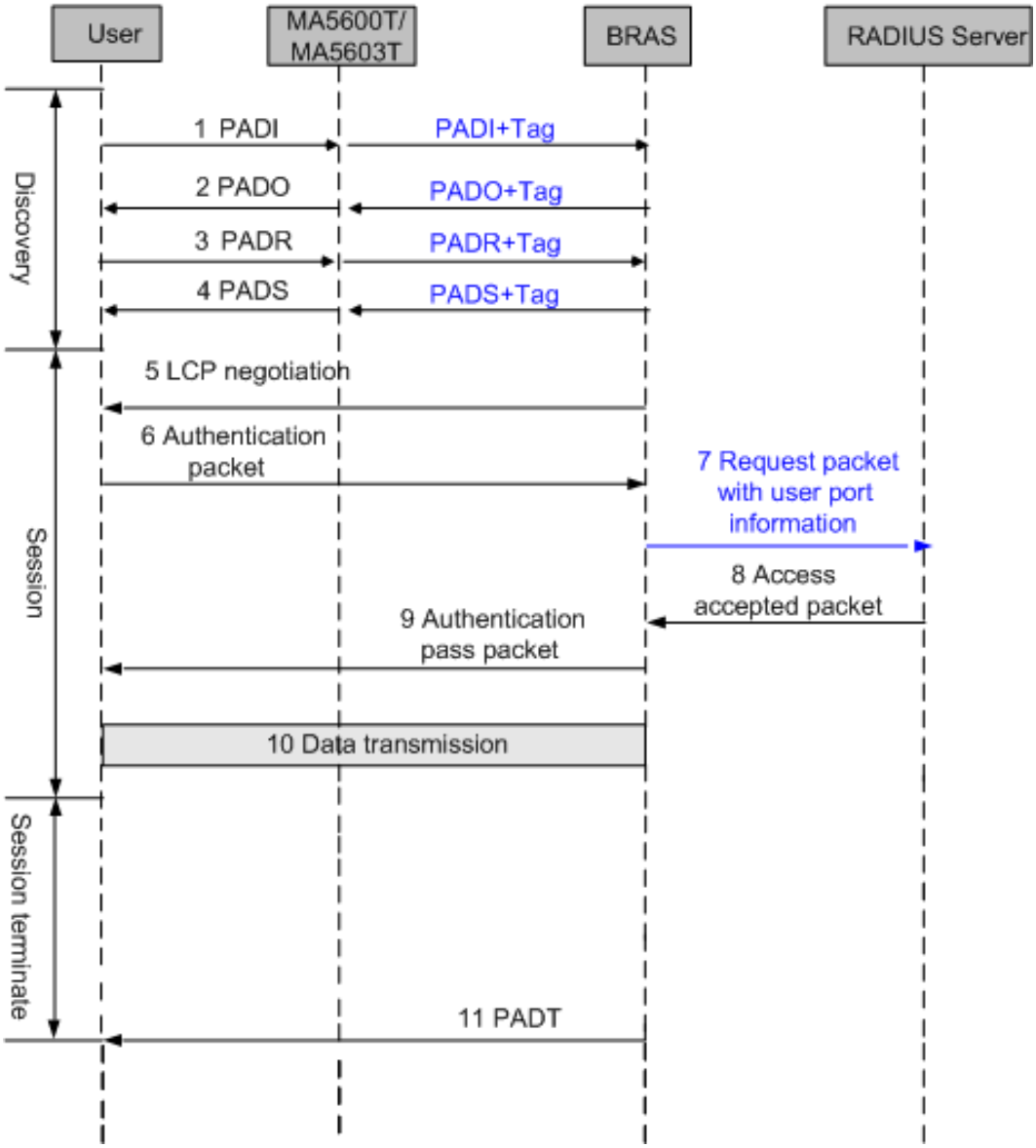
Principles

The Policy Information Transfer Protocol (PITP) can be enabled or disabled at the global, port, service port, or VLAN level. It takes effect only after being enabled at the four levels. Then, an access device can send the information about the physical ports of access users to the broadband remote access server (BRAS).

Working Principle of the PITP P Mode

Figure 27-13 shows the PPPoE dialup process in PITP P mode.

Figure 27-13 PPPoE dialup process in PITP P mode



NOTE

- PADI: refers to PPPoE active discovery initiation and is an initialization packet at the discovery stage.
- PADO: refers to PPPoE active discovery offer and is a response packet at the discovery stage.
- PADR: refers to PPPoE active discovery request and is a request packet at the discovery stage.
- PADS: refers to PPPoE active discovery session-confirmation and is a session confirmation packet at the discovery phase.
- PADT: refers to PPPoE active discovery terminate and is a session termination packet at the discovery phase.

The authentication in PITP P mode involves three stages: discovery, session, and session termination.

- Discovery stage: At this stage, the MA5600T/MA5603T/MA5608T adds a vendor tag to the upstream PADI and PADR packets and removes the vendor tag carried in the

downstream PADO and PADS packets. Then, the BRAS receives the packets carrying a vendor tag. By parsing the vendor tag, the BRAS obtains the information about the physical port of an access user.

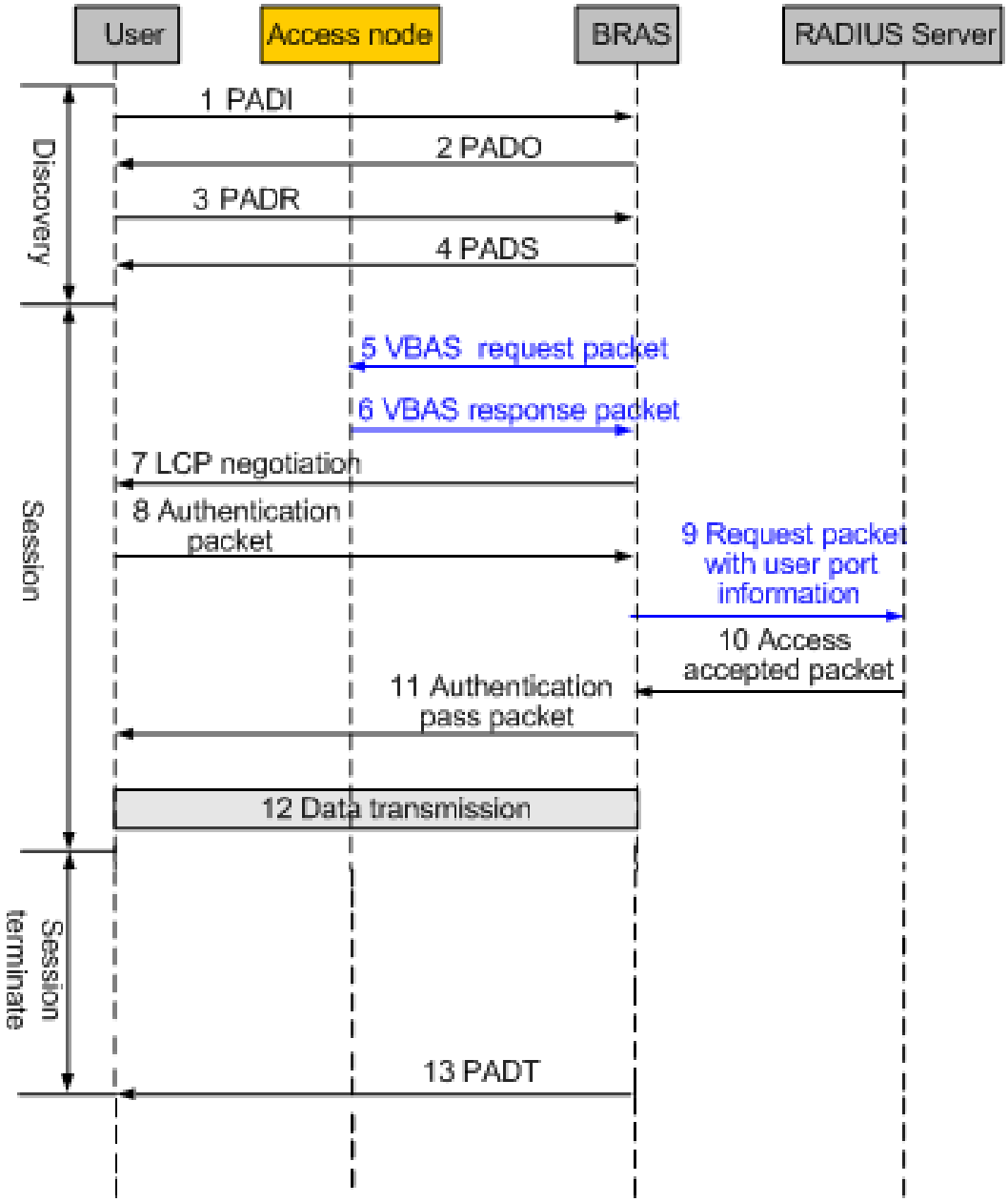
- Session stage: When sending a request packet to the RADIUS server, the BRAS provides the obtained information about the physical port as well as the user account and password to the RADIUS server. Based on the information, the RADIUS server determines whether to accept the access request. Specifically, if the user account matches the physical port, the RADIUS server accepts the access request. If the user account does not match the physical port, the RADIUS server rejects the access request. A user can transmit data only after being authenticated.
- Session termination stage: A PADT packet terminates a PPPoE session and can be sent at any time after a session is set up. It can be sent from the BRAS or an access user. The MA5600T/MA5603T/MA5608T (access device) does not add a vendor tag to a PADT packet.

In PITP P mode, an MA5600T/MA5603T/MA5608T adds the information about the physical port of an access user to the PPPoE packets at the PPPoE discovery stage. The physical port information can be used by the BRAS and RADIUS server to authenticate the access user. The other PPPoE dialup process in PITP P mode is the same as that when the PITP P mode is disabled.

Working Principle of the PITP V Mode

Figure 27-14 shows the PPPoE dialup process in PITP V mode.

Figure 27-14 PPPoE dialup process in PITP V mode



In PITP V mode, during the PPPoE negotiation between an access user and the BRAS, the BRAS proactively sends a VBRAS request packet to the MA5600T/MA5603T/MA5608T, asking the MA5600T/MA5603T/MA5608T to report the information about the physical port of the access user. The MA5600T/MA5603T/MA5608T sends the BRAS a VBAS response packet that carries the physical port information.

The authentication in PITP V mode involves three stages: discovery, session, and session termination.

- Discovery stage: The packet processing remains the same, regardless of whether the virtual broadband access server (VBAS) function is enabled. After the discovery stage ends, the BRAS sends a VBAS request packet to the MA5600T/MA5603T/MA5608T. Then, the MA5600T/MA5603T/MA5608T queries the physical port information

according to the VBAS request packet and sends a VBAS response packet carrying the physical port information to the BRAS. The BRAS parses the VBAS response packet and obtains the physical port information.

- Session stage: When sending a request packet to the RADIUS server, the BRAS provides the obtained information about the physical port as well as the user account and password to the RADIUS server. Based on the information, the RADIUS server determines whether to accept the access request. Specifically, if the user account matches the physical port, the RADIUS server accepts the access request. If the user account does not match the physical port, the RADIUS server rejects the access request. A user can transmit data only after being authenticated.
- Session termination stage: A PADT packet terminates a PPPoE session and can be sent at any time after a session is set up. It can be sent from the BRAS or an access user. The MA5600T/MA5603T/MA5608T (access device) does not add a vendor tag to a PADT packet.

PITP Packet Processing Policy Used by the MA5600T/MA5603T/MA5608T

Run the **pitp permit-forwarding service-port** command to configure whether a service port allows upstream packets to carry a vendor tag. Table 27-9 lists the PITP packet processing policy used by the MA5600T/MA5603T/MA5608T.

Table 27-9 PITP packet processing policy used by the MA5600T/MA5603T/MA5608T

Status of the PITP Function in Global or VLAN Service Profile Mode	Upstream Packets Carrying a Vendor Tag on a Service Port	Whether PITP Packets Carry a Vendor Tag	PITP Packet Processing Policy Used by the MA5600T/MA5603T/MA5608T
enable	**enable**	Yes	The MA5600T/MA5603T/MA5608T replaces the vendor tag carried in upstream PITP packets with the local one.
		No	The MA5600T/MA5603T/MA5608T adds the local vendor tag to upstream PITP packets.
	disable	Yes	The MA5600T/MA5603T/MA5608T discards upstream PITP packets.
		No	The MA5600T/MA5603T/MA5608T adds the local vendor tag to upstream PITP packets.
disable	**enable** or **disable**	Yes or No	The MA5600T/MA5603T/MA5608T forwards user-side PITP packets.
forward **NOTE** **forward** is supported	**enable**	Yes	The MA5600T/MA5603T/MA5608T forwards upstream PITP packets.
		No	The MA5600T/MA5603T/MA5608T adds the local vendor tag to upstream

Status of the PITP Function in Global or VLAN Service Profile Mode	Upstream Packets Carrying a Vendor Tag on a Service Port	Whether PITP Packets Carry a Vendor Tag	PITP Packet Processing Policy Used by the MA5600T/MA5603T/MA5608T
only in PITP P mode.			PITP packets.
	disable	Yes	The MA5600T/MA5603T/MA5608T discards upstream PITP packets.
		No	The MA5600T/MA5603T/MA5608T adds the local vendor tag to upstream PITP packets.
rebuild **NOTE** **rebuild** is supported only in PITP P mode.	**enable**	Yes	The MA5600T/MA5603T/MA5608T rebuilds the vendor tag by adding the local vendor tag to the original one.
		No	The MA5600T/MA5603T/MA5608T adds the local vendor tag to upstream PITP packets.
	disable	Yes	The MA5600T/MA5603T/MA5608T discards upstream PITP packets.
		No	The MA5600T/MA5603T/MA5608T adds the local vendor tag to upstream PITP packets.

Pay attention to the following points when using PITP on a network that contains an optical line terminal (OLT) and an optical network unit (ONU):

1. If PITP is enabled only on the OLT, PITP packets carry only the information about PON ports on the OLT.
2. If PITP is enabled only on the ONU, PITP packets carry only the information about user ports on the ONU.
3. If PITP is enabled on the OLT and ONU, run the **pitp permit-forwarding service-port** command on the OLT to configure whether a service port allows upstream packets to carry a vendor tag.
   - If the service port allows upstream packets to carry a vendor tag, PITP packets carry only the information about PON ports on the OLT.
   - If the service port does not allow upstream packets to carry a vendor tag, the dialing service is unavailable for the user of the service port because PADI packets cannot be sent in PITP P mode.

Generally, enable PITP on the OLT in global mode. To identify each ONU user, enable PITP on the ONU. The reason is that some PON ports on the OLT connect to ONUs and an ONU may connect to multiple users on a network, such as a fiber to the building (FTTB) network. In this case, an ONU is a user of the OLT.

General Packet Format in PITP P Mode

Figure 27-15 shows the general format of a PPPoE packet carrying a vendor tag in PITP P mode.

Figure 27-15 General format of a PPPoE packet carrying a vendor tag in PITP P mode

In the preceding figure:

- 0x0105 indicates a vendor tag.
- VendorID contains the sub-tags of the circuit ID (CID), remote ID (RID), sub-options, and IWF. VendorID has a fixed value of 0x00000DE9. In addition, VendorID in the vendor tag of an upstream packet is verified in PITP P mode.
- IWF identifies a PPPoA-to-PPPoE upstream packet.

Based on the vendor tag, the BRAS and RADIUS server implements a control policy.

NOTE
The specific format of a packet providing the information about the physical port of an access user to the BRAS is determined by the relay agent info option (RAIO) mode. For details about specific packet formats in PITP P mode, see 27.8.1 RAIO.

General Packet Format in PITP V Mode

The PITP V mode uses the VBAS protocol, which is a non-standard protocol proposed by China Telecom. The user information carried in VBAS packets is displayed as INFO in RAIO packets. Figure 27-16 shows the general format of a PPPoE packet in PITP V mode.

NOTE
INFO: indicates information, which is dedicated for working objects in PITP V mode and filled with a CID.

Figure 27-16 General format of a PPPoE packet in PITP V mode

NOTE
The specific format of a packet providing the information about the physical port of an access user to the BRAS is determined by the RAIO mode. For details about specific packet formats in PITP V mode, see 27.8.1 RAIO.

Configuring Anti-theft and Roaming of User Accounts Using PITP

Application Context

The Policy Information Transfer Protocol (PITP) is used to provide the information about the physical ports of access users to the broadband remote access server (BRAS). After obtaining the physical port information, the BRAS authenticates the binding between the user accounts and the physical ports, thereby preventing user account theft and roaming. PITP is available in

the PITP P (PPPoE+) mode and PITP V (VBAS) mode. VBAS refers to virtual broadband access server.

PITP applies to a network where an MA5600T/MA5603T/MA5608T works as an independent NE or is cascaded.

- When the MA5600T/MA5603T/MA5608T works as an independent NE, PC 1 and PC 2 connect to different ports on the MA5600T/MA5603T/MA5608T to perform PPPoE dialup to access the Internet.
- When the MA5600T/MA5603T/MA5608T is cascaded, PC 1 connects to the MA5600T/MA5603T/MA5608T and PC 2 connects to the device cascaded with the MA5600T/MA5603T/MA5608T to perform PPPoE dialup to access the Internet.

The working principles are the same in the two scenarios. PC 1 uses its account to perform PPPoE dialup to access the Internet. The BRAS binds the account of PC 1 to the physical location of PC 1 reported by the MA5600T/MA5603T/MA5608T. When PC 2 performs PPPoE dialup to access the Internet, if PC 2 forges the account of PC 1, the BRAS determines that the account of PC 2 does not match its physical location and rejects PC 2 to access the Internet.

Default Configuration

Table 27-10 lists the default PITP settings.

Table 27-10 Default PITP settings

Parameter	Default Setting
PITP function	Global status: disabled Port-level status: enabled VLAN-level status: enabled Service port-level status: enabled
PITP sub-option 90 function	Disabled
Whether a service port allows user-side PPPoE packets to carry a vendor tag	No

Context

The format of a packet providing the physical port of an access user to the BRAS is determined by the relay agent info option (RAIO) mode. Therefore, configure a RAIO mode before configuring PITP.

A RAIO mode can be configured in global or profile mode. A RAIO mode configured in profile mode takes precedence over that in global mode.

Procedure

Configure a RAIO mode in global mode.

1. Run the **raio-mode** ***mode*** **pitp-pmode** or **raio-mode** ***mode*** **pitp-vmode** command to configure the RAIO mode in PITP P mode or PITP V mode.

NOTE

The PITP P mode supports all RAIO modes. The PITP V mode supports only the device pre-defined and user-defined modes. If the type of the configured service flow is autosensing, the VPI/VCI in the tag is filled with 8191/35, regardless of whether the service flow has learned the VPI/VCI. VPI refers to virtual path identifier, and VCI refers to virtual channel identifier.

2. (Optional) If the RAIO mode is used-defined, run the **raio-format pitp-pmode** or **raio-format pitp-vmode** command to configure the RAIO mode in PITP P mode or PITP V mode.
   - In a user-defined RAIO mode, configure the circuit ID (CID) and remote ID (RID). The formats of the CID and RID are the same in PITP V mode.
   - In a non-user-defined RAIO mode, the tag formats are fixed and do not need to be configured.
   - If an access mode is specified, the configured tag formats take effect only in this access mode. If no access mode is specified, the configured tag formats take effect in all access modes.

NOTE

For details about tag formats, see the **raio-format** command.

Step 1 Configure a RAIO mode in profile mode.

1. Run the **raio-profile** command to create a RAIO profile.
2. Run the **raio-mode** ***mode*** **pitp-pmode** or **raio-mode** ***mode*** **pitp-vmode** command to configure the RAIO mode in PITP P mode or PITP V mode.

NOTE

The PITP P mode supports all RAIO modes. The PITP V mode supports only the device pre-defined and user-defined modes. If the type of the configured service flow is autosensing, the VPI/VCI in the tag is filled with 8191/35, regardless of whether the service flow has learned the VPI/VCI.

3. (Optional) If the RAIO mode is used-defined, run the **raio-format pitp-pmode** or **raio-format pitp-vmode** command to configure the RAIO mode in PITP P mode or PITP V mode.
   - In a user-defined RAIO mode, configure the CID and RID. The formats of the CID and RID are the same in PITP V mode.
   - In a non-user-defined RAIO mode, the tag formats are fixed and do not need to be configured.
   - If an access mode is specified, the configured tag formats take effect only in this access mode. If no access mode is specified, the configured tag formats take effect in all access modes.

NOTE

For details about the RAIO format, see the **raio-format** command.

4. Run the **quit** command to quit the RAIO profile mode.
5. Run the **vlan bind raio-profile** command to bind the RAIO profile created in Step 2.1 to a VLAN.

Step 2 Enable the PITP function.

The PITP function can be enabled or disabled at four levels: global-level, port-level, VLAN-level, and service port-level. It takes effect only after being enabled at the four levels.

1. Run the **pitp enable pmode** command to enable the PITP P mode globally.

   In PITP V mode, run the **pitp vmode ether-type** command to configure an Ethernet protocol type on the MA5600T/MA5603T/MA5608T and ensure that the Ethernet protocol type on the MA5600T/MA5603T/MA5608T is the same as that on the BRAS. Then, run the **pitp enable vmode** command to enable the PITP V mode globally.

NOTE
Configure an Ethernet protocol type when the PITP V mode is disabled.

2. Run the **pitp port** or **pitp board** command to enable the PITP function at the port level.
3. Enable the PITP function at the VLAN level.
   a. Run the **vlan service-profile** command to create a VLAN service profile.
   b. Run the **pitp enable** command to enable the PITP function at the VLAN level.
   c. Run the **commit** command to make the profile configuration take effect.
   d. Run the **quit** command to quit the VLAN service profile mode.
   e. Run the **vlan bind service-profile** command to bind the VLAN service profile created in Step 3.3.a to a VLAN.
4. Run the **pitp service-port** command to enable the PITP function at the service port level.

Step 3 Configure optional PITP attributes.

1. Run the **pitp permit-forwarding service-port** command to configure a service port to allow user-side PPPoE packets to carry a vendor tag.
2. In PITP P mode, run the **pitp sub-option90** command to enable the sub-option 90 function.

   Only the PITP P mode supports the reporting of sub-option 90's line parameters, including activated bandwidths.

----End

Result

After the configuration, you can access the Internet through PPPoE dialup.

Example

The following configurations are used as an example to enable the PITP P mode on service port 1 on an 0/2/0:

- RAIO mode: user-defined
- ATM access mode:
  a. CID format: atm
  b. Subrack ID/slot ID/port ID: vpi.vci
- Ethernet access mode:
  a. CID format: eth
  b. Subrack ID/slot ID/port ID: vlanid
- xPON access mode:
  a. CID: xpon
  b. Subrack ID/slot ID/port ID: ontid.vlanid
  c. RID format: labels of service ports

```
huawei(config)#raio-mode user-defined pitp-pmode
huawei(config)#raio-format pitp-pmode cid atm anid atm frame/slot/port:vpi.vci
huawei(config)#raio-format pitp-pmode cid eth anid eth frame/slot/port:vlanid
huawei(config)#raio-format pitp-pmode cid xpon anid xpon frame/slot/port:ontid.vlanid
huawei(config)#raio-format pitp-pmode rid atm plabel
huawei(config)#raio-format pitp-pmode rid eth plabel
```

```
huawei(config)#raio-format pitp-pmode rid xpon plabel
huawei(config)#pitp enable pmode
huawei(config)#pitp port 0/2/0 enable
huawei(config)#pitp service-port 1 enable
```

The following configurations are used as an example to configure the Ethernet protocol type of VBRAS packets to be the same as that of the upper-layer BRAS (0x8500) and enable the PITP V mode at the port, service port, and VLAN levels on service port 0:

- RAIO profile index: 10
- RAIO mode: user-defined
- ATM access mode:
  a. CID format: atm
  b. Subrack ID/slot ID/port ID: vpi.vci
- Ethernet access mode:
  a. CID or RID format: eth
  b. Subrack ID/slot ID/port ID: vlanid
- xPON access mode:
  a. CID or RID format: xpon
  b. Subrack ID/slot ID/port ID: ontid.vlanid
- ID of the VLAN for service port 0: 11

```
huawei(config)#raio-profile index 10
huawei(config-raio-profile-10)#raio-mode user-defined pitp-vmode
huawei(config-raio-profile-10)#raio-format pitp-vmode atm anid atm
frame/slot/port:vpi.vci
huawei(config-raio-profile-10)#raio-format pitp-vmode eth anid eth
frame/slot/port:vlanid
huawei(config-raio-profile-10)#raio-format pitp-vmode xpon anid xpon
frame/slot/port:ontid.vlanid
huawei(config-raio-profile-10)#quit
huawei(config)#vlan bind raio-profile 11 index 10
huawei(config)#pitp vmode ether-type 0x8500
huawei(config)#pitp enable vmode
huawei(config)#pitp port 0/2/0 enable
huawei(config)#pitp service-port 0 enable
```

27.9 ARP/NS Security

27.9.1 Introduction

ARP/NS Proxy

After ARP/NS proxy reply is enabled, the system searches for user's going online information based on the destination IP address and VLAN after receiving ARP request (broadcast) or NS multicast packets from the network side. If a user goes online, the system performs proxy reply (not forwarding the ARP request or NS multicast packets to user side). If no user goes

online, the system discards or forwards the ARP request or NS multicast packets according to the configure policy by operators.

ARP/NS proxy reply avoids sending ARP or NS multicast packets to irrelevant users, improving system security.

NOTE

ARP broadcast packets and NS multicast packets sent to irrelevant users allow a malicious user can get IP of the normal user, may causing the attack to normal user.

ARP Broadcast or NS Multicast Converted to Unicast

MA5600T/MA5603T/MA5608T supports the ARP broadcast or NS multicast converted to unicast function on the network side, and the configuration of the forwarding policy after the ARP broadcast or NS multicast converted to unicast.

27.9.2 Principle

ARP Proxy Reply

After enabled the ARP proxy reply function by **security arp-reply** command, the MA5600T/MA5603T/MA5608T searches for user's going online information based on the destination IP address (IPv4 or IPv6 address) and VLAN after it receives network-side ARP broadcast request packets. If a user goes online, the MA5600T/MA5603T/MA5608T performs proxy reply (not forwarding ARP request packets to user side). If no user goes online, the MA5600T/MA5603T/MA5608T process the ARP request packets according to the **security arp-reply unknown-policy** command configuration as follows.

- Setting the policy to **discard**, the MA5600T/MA5603T/MA5608T broadcasts the ARP request packets to cascading-side and network-side ports excluding the source port in the VLAN. The user side does not receive the network-side ARP request packets.
- Setting the policy to **forward**, the MA5600T/MA5603T/MA5608T broadcasts the ARP request packets to user-side, cascading-side, and network-side ports excluding the source port in the VLAN.

NOTE

- By default, the forwarding policy of ARP broadcast request packets is forward when no user goes online.
- Unicast ARP request packets are not effected by ARP proxy reply function.
- Broadcasts the ARP packet if this packet is a gratuitous ARP packet (such a packet is used for address announcement, not for address resolution).

NS Proxy Reply

After enabled the Neighbor Solicitation (NS) proxy reply function by **security ns-reply** command, the MA5600T/MA5603T/MA5608T searches for user's going online information based on the destination IPv6 address and VLAN after it receives network-side NS packets. If a user goes online, the MA5600T/MA5603T/MA5608T performs proxy reply (not forwarding NS packets to user side). If no user goes online, the MA5600T/MA5603T/MA5608T process the NS packets according to the **security ns-reply unknown-policy** command configuration as follows.

- Setting the policy to **discard**, the MA5600T/MA5603T/MA5608T broadcasts the NS packets to cascading-side and network-side ports excluding the source port in the VLAN. The user side does not receive the network-side ARP request packets.

- Setting the policy to **forward**, the MA5600T/MA5603T/MA5608T broadcasts the NS packets to user-side, cascading-side, and network-side ports excluding the source port in the VLAN.

NOTE
By default, the forwarding policy of NS packets is forward when no user goes online.

ARP Broadcast Converted to Unicast

MA5600T/MA5603T/MA5608T supports the ARP broadcast converted to unicast function on the network side and the forwarding policy configuration.

If the ARP broadcast converted to unicast function is enabled on the network side, MA5600T/MA5603T/MA5608T queries user's going online information based on destination IP address and VLAN after receiving an ARP request.

- After detecting that a user goes online, MA5600T/MA5603T/MA5608T sends the ARP request to this user.
- If no online users are detected, when **security arp-unicast unknown-policy** is set to **discard**, the ARP request is broadcast cascading-side and network-side ports excluding the source port in the VLAN. However, the user side cannot receive ARP requests sent by the network side; when **security arp-unicast unknown-policy** is set to **forward**, the ARP request is broadcast to the user-side, cascading-side, and network-side ports excluding the source port in the VLAN.

NS Multicast Converted to Unicast

MA5600T/MA5603T/MA5608T supports the NS multicast converted to unicast function on the network side.

If the NS multicast converted to unicast function is enabled on the network side, MA5600T/MA5603T/MA5608T queries user's going online information based on destination IP address and VLAN after receiving a NS request.

- After detecting that a user goes online, MA5600T/MA5603T/MA5608T sends the NS request to this user.
- If no online users are detected, when **security arp-unicast unknown-policy** is set to **discard**, the NS request is broadcast to cascading-side and network-side ports excluding the source port in the VLAN. However, the user side cannot receive NS requests sent by the network side; when **security arp-unicast unknown-policy** is set to **forward**, the NS request is broadcast to the user-side, cascading-side, and network-side ports excluding the source port in the VLAN.

27.9.3 Feature Updates

Table 27-11 Updates in the ARP/NS proxy reply feature

Product	Version
V800R015C00	The ARP broadcast or NS multicast converted to unicast function on the network side is supported.
V800R009C00	It is the first version that supports this feature.

28 MAC Address Security Features

About This Chapter

The access node provides multiple MAC address security features to protect networks against forged MAC addresses. A combined use of these features can protect networks against malicious attacks in various scenarios.

28.1 MAC Address Security Threats

When the access node works in VLAN+MAC forwarding mode, the access node faces three common MAC address security threats: user MAC address spoofing, upper-layer device MAC address spoofing, and MAC address exhaustion.

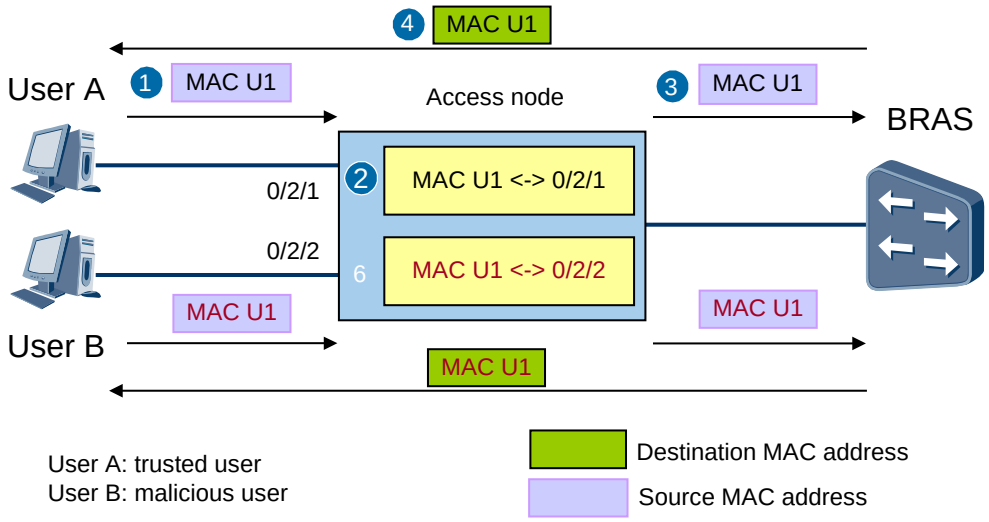
The access node supports multiple forwarding modes. When the access node works in VLAN+MAC forwarding mode (that is, the VLAN forwarding mode is set to **vlan-mac**), packets are forwarded based on VLANs and MAC addresses. VLANs can be created only manually. Therefore, it is difficult for malicious users to tamper with the VLAN configuration. The MAC address entries on the access node can be manually configured or dynamically learned by the access node. The manually configured MAC address entries are static MAC address entries, while the MAC address entries learned by the access node are dynamic MAC address entries. It is difficult for malicious users to tamper with the static MAC address entries. Therefore, malicious users mainly use the dynamic MAC address entries to attack the network. In the VLAN+MAC forwarding mode, the dynamic MAC address learning mechanism enables malicious users to attack the network by forging the source MAC addresses of trusted users or upper-layer devices. This is a most common attack method. Common MAC address security threats are as follows:

- User MAC address spoofing
- Upper-layer device MAC address spoofing
- MAC address exhaustion

User MAC Address Spoofing

A malicious user accesses a network by forging the source MAC address of a trusted user. The malicious user occupies the network resources of the trusted user, resulting in service interruption of the trusted user. The following uses Internet access in PPPoE dialup mode as an example. Figure 28-1 shows the principle of user MAC address spoofing.

Figure 28-1 User MAC address spoofing



The following details the process of user MAC address spoofing:

1. User A sends PPPoE dialup packets. The source MAC address carried in the packets is MAC U1.
2. The access node learns the mapping between the source MAC address and the port (MAC U1 <-> 0/2/1) from the packets sent by user A, and records the mapping in the MAC address table.
3. The access node forwards the packets sent by user A to the broadband remote access server (BRAS).
4. After authenticating the packets, the BRAS sends packets using MAC U1 as the destination MAC address to the access node. Then the access node forwards the packets to user A through port 0/2/1 based on the mapping between the MAC address and the port recorded in the MAC address table. After receiving the packets, user A can access the network.
5. User B also sends PPPoE dialup packets. The packets carry a forged source MAC address (that is, the source MAC address of user A, MAC U1).
6. The access node learns the mapping between the source MAC address and the port (MAC U1 <-> 0/2/2) from the packets sent by user B, and records the mapping in the MAC address table. As a result, the port corresponding to MAC U1 is changed from 0/2/1 to 0/2/2 in the MAC address table.
7. The access node forwards the packets sent by user B to the BRAS.
8. After authenticating the packets, the BRAS sends packets using MAC U1 as the destination MAC address to the access node. Then the access node forwards the packets to user B through port 0/2/2 based on the mapping between the MAC address and the port updated in the MAC address table.

In this case, user B accesses the network by forging the source MAC address of user A. The packets that should be sent to user A are forwarded to user B, which enables user B to intercept communication data of user A. As a result, user B occupies the communication resources of user A and communication of user A is interrupted.

Upper-Layer Device MAC Address Spoofing

Malicious users forge the MAC address of an upper-layer device to intercept the communication data forwarded to the upper-layer device. The principle of spoofing upper-layer device MAC addresses is as follows:

1. A malicious user sends packets to the access node using the MAC address of the upper-layer device as the source MAC address.
2. After receiving the packets sent by the malicious user, the access node learns the MAC address and updates its MAC address table using the learned MAC address. In other words, the access node learns the MAC address of the upper-layer device from the port of the malicious user other than from the upstream port.
3. The access node forwards packets sent by other users to the port of the malicious user instead of to the upper-layer device.

When 13.7 Layer 2 User Bridging is disabled, user ports are isolated from each other at Layer 2 and Layer 2 forwarding cannot be implemented. Therefore, the packets sent by other users to the upper-layer device are discarded, resulting in communication interruption.

MAC Address Exhaustion

Malicious users attack the access node and impair network communication by sending a large number of packets with different forged source MAC addresses. The principle of exhausting MAC address resources is as follows:

1. A malicious user forges a large number of different source MAC addresses.
2. The access node learns a large number of junk entries from the packets sent by the malicious user, which consume resources in the MAC address table.
3. After MAC address table resources are exhausted, the access node cannot learn new MAC addresses. As a result, packets sent by new users can only be forwarded as unknown unicast packets.
4. Based on the forwarding policy for unknown unicast packets, the user packets are broadcast or discarded.
   - When the unknown unicast traffic suppression function is enabled on the access node, the user packets are discarded and therefore the new users fail to access the network.
   - When the unknown unicast traffic suppression function is disabled on the access node, the user packets are broadcast, which consumes bandwidth and affects network communication.

28.2 MAC Address Security Solutions

This topic describes MAC address security features and multiple combinations of the security features that can safeguard the access node in various scenarios.

MAC Address Security Features

To address MAC address security issues, the access node supports multiple security features, as shown in Table 28-1.

Table 28-1 MAC address security features

Security Feature	Description
28.3 MAC Anti-Spoofing	This feature is implemented by binding dynamic source MAC addresses to service streams and filtering dynamic source MAC addresses. The access node monitors the interaction process of PPPoE, DHCP, DHCPv6, and StateLess Address Auto Configuration (SLAAC) protocol packets, and dynamically generates MAC address binding entries and filtering entries. Then the access node filters MAC address spoofing packets based on the entries, preventing malicious users from forging the source MAC addresses of other users or upper-layer devices.
28.4 Static MAC Address Binding	This feature is implemented by binding the static source MAC address of a user to a service stream, protecting the user's MAC address from forgery and preventing the user from forging the source MAC addresses of other users or upper-layer devices.
28.5 Static MAC Address Filtering	This feature is implemented by setting the MAC address of an upper-layer device to the source MAC address to be filtered, preventing the MAC address of the upper-layer device from being used by malicious users.
28.6 MAC Anti-Duplicate	Before the source MAC address of a port is aged, the access node learns the source MAC address of the port from another port and then updates the mapping between the source MAC address and the port in the MAC address table. This process is called MAC address duplicate (also known as MAC address flapping), because it can be regarded as if the access node copied the MAC address from one port to another port (the MAC address flaps from one port to another port). The access node supports the MAC anti-duplicate feature to prevent the MAC addresses of trusted users or upper-layer devices from being forged by malicious users.
28.7 VMAC	The access node replaces the source MAC addresses of users with virtual MAC addresses, ensuring uniqueness of MAC addresses on the entire network and preventing MAC address conflicts. This feature also prevents untrusted MAC addresses from accessing carriers' networks and prevents malicious users from forging the source MAC addresses of trusted users or upper-layer devices.

MAC Address Security Solutions

The user can adopt different combinations of the above-mentioned features to address three MAC address security issues. Table 28-2 lists the security solutions recommended in different scenarios.

Table 28-2 MAC address security solutions

Solution	User MAC Address Spoofing	Upper-Layer Device MAC Address Spoofing	MAC Address Exhaustion	Application Scenario and Limitation
Solution 1	28.3 MAC Anti-Spoofing	28.3 MAC Anti-Spoofing	The maximum number of MAC addresses dynamically bound to a service stream is limited.	Optimal choice for PPPoE, DHCP, DHCPv6, and SLAAC users. This solution has the following advantages: • The configuration is simple and flexible. • Statistics collection, logging, and alarming functions are supported, facilitating location of malicious users. However, this solution does not work if malicious users go online before trusted users.
Solution 2	28.4 Static MAC Address Binding	28.5 Static MAC Address Filtering	N/A	This solution is mainly used by static private line users whose IP addresses are manually configured. The MAC addresses of users and upper-layer devices should be obtained before configuring static source MAC address binding and static source MAC address filtering on the access node. After the configuration is complete, the MAC addresses of users and upper-layer devices cannot be changed. Therefore, configuration workload is heavy and the configuration is not flexible.
Solution 3	28.6 MAC Anti-Duplicate	28.6 MAC Anti-Duplicate	The maximum number of dynamic MAC addresses that can be learned is limited.	Only some hardware supports the MAC anti-duplicate function. Therefore, this solution applies to scenarios with such hardware. This solution has the following limitations: • The MAC anti-duplicate function cannot be enabled or disabled based on VLANs. For example, if the carrier requires that the MAC anti-duplicate function be disabled on some VLANs (for example, the VLAN of private line services), this solution cannot apply. • Only the source MAC addresses

Solution	User MAC Address Spoofing	Upper-Layer Device MAC Address Spoofing	MAC Address Exhaustion	Application Scenario and Limitation
				in Ethernet packet headers are checked, but the user MAC addresses in packet payload are not checked. Therefore, this solution cannot prevent the source MAC addresses carried in DHCP packet payload from being forged. To address this issue, bind dynamic MAC addresses to service streams.
Solution 4	28.7.3 N:1 VMAC Principles	28.7.3 N:1 VMAC Principles	The maximum number of virtual MAC (VMAC) addresses to be allocated is limited.	This solution applies to PPPoA and PPPoE users but not IPoE users (including DHCP and static IP users). This solution has the following advantages: • A smaller number of user MAC addresses are used, occupying less MAC address resources on the access node. • Less resources in the MAC address table of upper-layer devices are occupied.
Solution 5	28.7.2 1:1 VMAC Principles	28.7.2 1:1 VMAC Principles	The maximum number of VMAC addresses to be allocated is limited.	This solution does not depend on the dialup process. Therefore, it applies to static and dynamic users. This solution supports easy operation and maintenance (O&M) because it enables upper-layer devices to directly locate user ports based on the VMAC addresses carried in packets.

28.3 MAC Anti-Spoofing

When users dynamically obtain IP addresses in PPPoE, DHCP, DHCPv6, or StateLess Address AutoConfiguration (SLAAC) dialup mode, the most frequently used MAC address security method is MAC anti-spoofing.

28.3.1 Introduction

MAC spoofing is a process in which malicious users impair network communication by forging the source MAC addresses of authorized users or upper-layer devices. For security threats resulting from MAC spoofing, see 28.1 MAC Address Security Threats.

When users dynamically obtain IP addresses in PPPoE, DHCP, DHCPv6, or StateLess Address AutoConfiguration (SLAAC) dialup mode, the most frequently used MAC address security method is MAC anti-spoofing. The MAC anti-spoofing feature includes dynamic source MAC address binding and dynamic source MAC address filtering, which protect networks against security threats listed in Table 28-3. After the MAC anti-spoofing feature is enabled, the dynamic source MAC address binding and dynamic source MAC address filtering functions are enabled.

Table 28-3 Two functions of the MAC anti-spoofing feature

Function	Security Threat
Dynamic source MAC address binding	• User MAC spoofing • MAC address exhaustion
Dynamic source MAC address filtering	Upper-layer device MAC spoofing

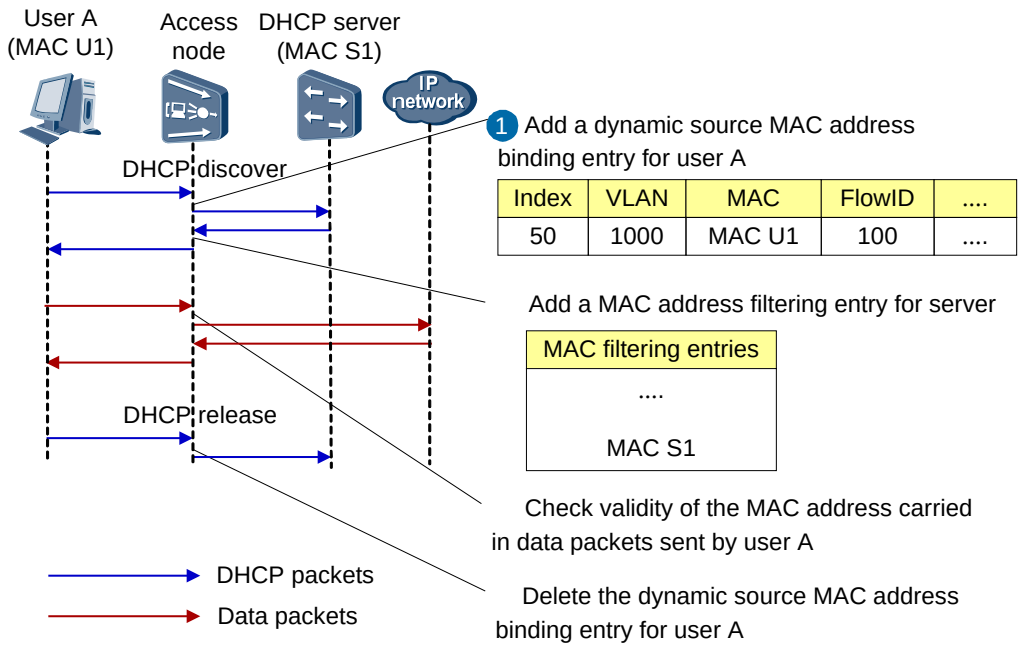
28.3.2 Principle

The MAC anti-spoofing feature consists of the dynamic source MAC address binding and dynamic source MAC address filtering functions. The following first uses DHCP users as an example to describe the process how users go online/offline after MAC anti-spoofing is enabled, and then describes principles of dynamic source MAC address binding and dynamic source MAC address filtering.

NOTE
In the following description, MAC addresses refer to source MAC addresses other than destination MAC addresses.

The access node automatically generates MAC address entries for users and servers by monitoring the interaction process of PPPoE, DHCP, DHCPv6, and StateLess Address AutoConfiguration (SLAAC) protocol packets, and then forwards or discards packets received through user ports based on the MAC address entries. Figure 28-2 shows the online/offline process of a DHCP user.

Figure 28-2 Online/Offline process of a DHCP user



1. When user A is getting online, the access node monitors the interaction process of DHCP packets between user A and the DHCP server to obtain the MAC address and IP address lease time of user A. On the access node, a dynamic source MAC address binding entry with index 50 is generated to record the VLAN, MAC address, and service port index (FlowID) of user A.
2. The access node learns the source MAC address (MAC S1) from the response packet sent by the DHCP server, and adds a MAC address filtering entry for MAC S1.
3. After user A goes online, the access node checks validity of the MAC address carried in data packets sent by user A based on the dynamic source MAC address binding entry. Only data packets using MAC U1 as the source MAC address can pass through the access node, and data packets using other source MAC addresses are discarded.
4. When detecting user A offline, the access node deletes the dynamic source MAC address binding entry for user A.

After MAC anti-spoofing is enabled, the access device will modify the exchange identification (XID) of the DHCP packet sent by the user, so that the XID of the DHCP packet sent by the DHCP client is different from that of the DHCP packet received by the DHCP server. Generally, the DHCP server does not verify the XID, and therefore services are not affected. If the carrier adds information into the XID of the packet sent by the DHCP client for DHCP server verification (this is not defined in the standard), the verification may fail and services will be affected.

NOTE

XID is a field carried by the DHCP packet, and it is defined by the standard. The XID is equivalent to the serial number of the DHCP packet.

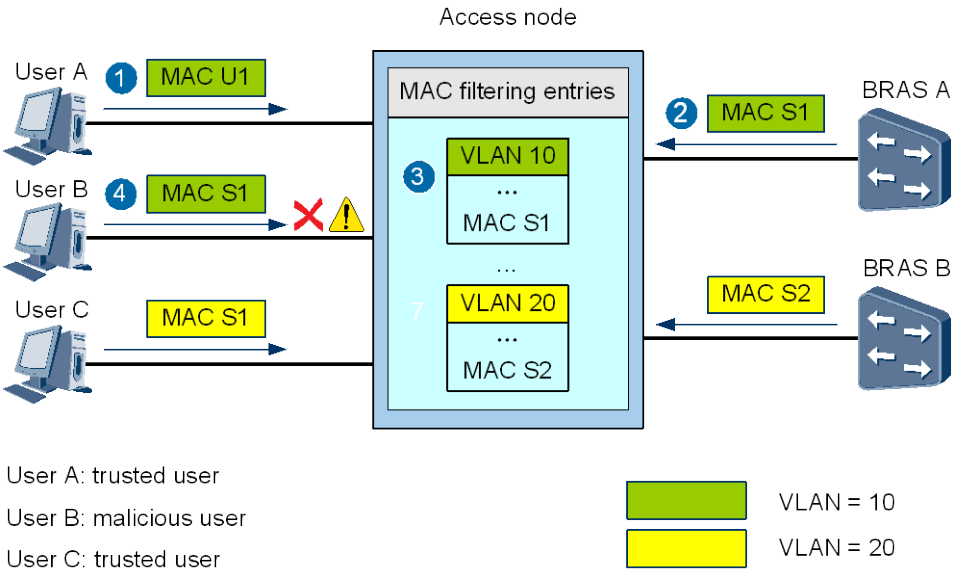
Dynamic Source MAC Address Filtering

The basic principle of MAC address filtering is that packets received through user ports are discarded if they carry MAC addresses that are the same as those in MAC address filtering entries. Dynamic source MAC address filtering is a process in which the access node

dynamically learns the MAC addresses, and adds the MAC addresses of upper-layer devices to MAC address filtering entries.

By monitoring the interaction process of protocol packets, the access node automatically adds the MAC addresses carried in packets sent by the server to MAC address filtering entries, and then filters packets based on the entries. The following uses Internet access in PPPoE dialup mode as an example. Figure 28-3 shows the principle of dynamic source MAC address filtering.

Figure 28-3 Dynamic source MAC address filtering



As shown in the preceding figure, the access node is connected to two broadband remote access servers (BRASs). User A, user B, and BRAS A are in VLAN 10, while user C and BRAS B are in VLAN 20. The following details the process of dynamic source MAC address filtering:

1. User A, a trusted user, sends PPPoE dialup packets to the access node. In the packets, the source MAC address is MAC U1 and the S-VLAN ID is 10.
2. After receiving the packets from user A, BRAS A sends response packets using MAC S1 as the source MAC address to user A.
3. The access node learns the source MAC address (MAC S1) from the packets sent by BRAS A, and adds the source MAC address to MAC address filtering entries of VLAN 10.
4. User B, a malicious user, sends packets using MAC S1 as the source MAC address to the access node. After receiving the packets, the access node finds that MAC S1 exists in the source MAC address filtering entries of VLAN 10 (where user B resides). Then the access node discards the packets and reports an event, preventing user B from attacking the network by forging the source MAC address of BRAS A.
5. User C, a trusted user, sends PPPoE dialup packets to the access node. In the packets, the source MAC address is MAC S1 and the S-VLAN ID is 20.
6. After receiving the packets from user C, BRAS B sends response packets using MAC S2 as the source MAC address to user C.

7. The access node learns the source MAC address (MAC S2) from the packets sent by BRAS B, and adds the source MAC address to MAC address filtering entries of VLAN 20.

The preceding process shows that dynamic source MAC address filtering takes effect only in a VLAN. In other words, the source MAC addresses carried in user packets in a VLAN can be the same as those of servers in other VLANs.

Dynamic Source MAC Address Binding

Dynamic source MAC address binding uses the following functions to protect networks against user MAC spoofing and MAC address exhaustion.

Function	Description
MAC address conflict check	This function ensures that users can only use the MAC addresses that are not used by other users to send dialup packets. It prevents MAC address conflicts caused by inappropriate address planning, and prevents malicious users from forging the MAC addresses of trusted users as well.
MAC address quantity check	This function limits the number of MAC addresses that can be used by a service port, protecting networks against MAC address exhaustion.
MAC address validity check	This function ensures that users can only use the MAC addresses bound to the service ports to send data packets and non-dialup protocol packets, preventing users from forging MAC addresses of other users or upper-layer devices.

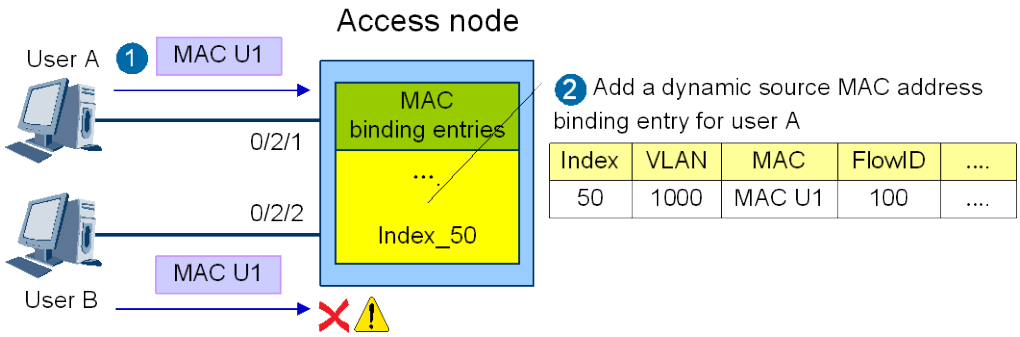
Only the MAC addresses carried in packets received through user ports are checked. The MAC addresses carried in packets received through upstream and cascading ports are not checked. This is because packets sent by upper-layer devices to upstream ports are generally trustable, and the MAC addresses carried in packets received through cascading ports have been checked on lower-layer devices.

MAC Address Conflict Check

The access node checks whether MAC address conflicts occur before generating new dynamic source MAC address binding entries. When the access node finds that the MAC addresses carried in dialup packets conflict with any one used by online users, the access node discards the packets and reports events. These discarded packets are called **MAC address conflicting packets**.

Figure 28-4 shows the principle of MAC address conflict check.

Figure 28-4 MAC address conflict check



The following details the process of MAC address conflict check:

1. User A sends dialup packets using MAC U1 as the source MAC address to the access node.
2. On the access node, a dynamic source MAC address binding entry with index 50 is generated to record the VLAN, MAC address, and service port index (FlowID) of user A. User A goes online after being authenticated by the server.
3. User B attempts to go online and sends dialup packets using MAC U1 as the source MAC address to the access node. After receiving the packets, the access node looks up MAC U1 in dynamic source MAC address binding entries. It finds that MAC U1 has been used by user A who is in the same VLAN as user B. The access node determines that an MAC address conflict occurs, and discards the packets of user B and reports an event. As a result, user B is prevented from going online and communication for user A is not affected.

If user A is a trusted user and user B is a malicious user, the preceding process can prevent the malicious user from forging the MAC address of the trusted user. However, if user A is a malicious user and user B is a trusted user, the preceding process cannot prevent MAC address forgery, because the dialup packets sent by the trusted user who attempts to go online later than the malicious user will be discarded as MAC address conflicting packets. As a result, the trusted user fails to go online.

When one of the following conditions is met, MAC address conflicts occur:

- New users' VLANs and MAC addresses are the same as those of other users in the dynamic source MAC address binding entries.
- New users' VLANs and MAC addresses are the same as the VLANs and static MAC addresses that have been configured for other service ports or physical ports.

MAC Address Quantity Check

The access node checks the number of bound MAC addresses before generating new dynamic source MAC address binding entries. Dialup packets of a new user will be discarded when one of the following conditions is met:

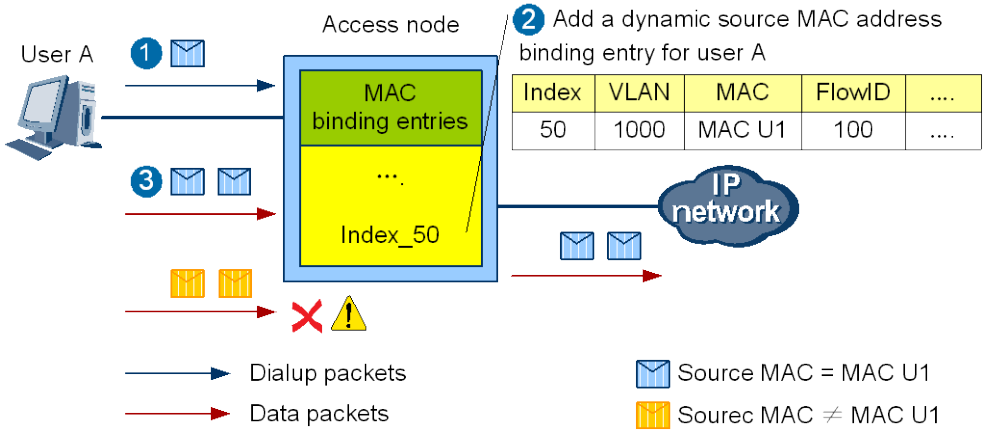
- In the dynamic source MAC address binding entries, the number of MAC addresses that have been bound to the service port reaches the maximum value.
- In the dynamic source MAC address binding entries, the number of MAC addresses that have been bound in a system reaches the maximum value.

MAC Address Validity Check

When receiving data packets and non-dialup packets, the access node checks whether the MAC addresses carried in the packets have been bound to service ports in the dynamic source MAC address binding entries. The access node regards the packets as valid only when the MAC addresses have been bound to service ports in the dynamic source MAC address binding entries. If the MAC addresses have not been bound, the access node discards the packets as **MAC spoofing packets** and reports events.

Figure 28-5 shows the principle of MAC address validity check.

Figure 28-5 MAC address validity check



The following details the process of MAC address validity check:

1. User A sends dialup packets using MAC U1 as the source MAC address to the access node.
2. After performing MAC address conflict check on the dialup packets sent by user A, the access node generates a dynamic source MAC address binding entry with index 50 to record the VLAN, MAC address, and service port index (FlowID) of user A. User A goes online after being authenticated by the server.
3. User A sends data packets using MAC U1 as the source MAC address to the access node.
4. The access node can find the MAC address in the dynamic source MAC address binding entries of user A. Therefore, the access node considers the data packets valid and forwards them to the upper-layer device.
5. User A sends data packets using a MAC address other than MAC U1 as the source MAC address to the access node. The access node cannot find the MAC address in the dynamic source MAC address binding entries of user A. Therefore, the access node regards the data packets as MAC spoofing packets, discards them, and reports an event.

The access node always checks MAC address validity on data packets. For protocol packets, the access node determines whether to check MAC address validity based on the following scenarios:

- Scenario 1: For dialup protocol packets that can trigger generation of dynamic source MAC address binding entries, the access node performs MAC address conflict check and MAC address quantity check but not MAC address validity check.

- Scenario 2: For Internet Group Management Protocol (IGMP) packets, the **security anti-macspoofing exclude** command can be run to enable or disable MAC address validity check.

User Going Offline

Users may go offline normally or abnormally. Table 28-4 lists normal offline conditions supported by different protocols.

Table 28-4 Normal offline conditions supported by different protocols

Normal Offline Condition	PPPoE	DHCP	DHCPv6	SLAAC
Leave packets sent by users or servers are received.	Yes	Yes	Yes	No
IP address lease time of users expires.	No	Yes	Yes	Yes

When MAC anti-spoofing is enabled, the access node can discover that users are offline by monitoring the interaction process of protocol packets if the users go offline normally. If the users go offline abnormally (for example, due to power-off of terminals), the access node cannot determine whether the users are online by monitoring the interaction process of protocol packets. To prevent offline users from occupying system resources, the access node provides the abnormal offline detection function.

- For PPPoE users, the access node periodically checks whether users go offline abnormally. If a user does not get online within a specified timeout period, the access node determines that the user goes offline abnormally. The check period and timeout time can be set by running the **security pppoe timeout** command.
- For DHCP, DHCPv6, and SLAAC users, the access node performs abnormal offline detection based on the conditions listed in Table 28-5.

Table 28-5 Abnormal offline detection for DHCP, DHCPv6, and SLAAC users

Trigger Condition	Monitoring Result	Operation
When receiving new dialup packets, the access node performs abnormal offline detection if one of the following conditions is met: • New users' VLANs and MAC addresses are the same as those of other users in the dynamic source MAC address binding entries. • The number of dynamic MAC addresses that have been bound to the service ports of new users reaches the maximum value. • The number of MAC addresses that have been bound in the system	Users are online.	The dialup packets are discarded.
	Users go offline abnormally.	The access node deletes the dynamic source MAC address binding entries of the users, and generates new dynamic source MAC address binding entries based on the new dialup packets.

Trigger Condition	Monitoring Result	Operation
reaches the maximum value.		

When users go offline abnormally, time difference exists between the time point at which the users go offline and the time point at which the access node detects the users offline.

Downstream PADT Packet Processing

In V800R013C00, the **security anti-macspoofing ignore** command is newly supported. Using this command, the system can be configured not to process certain types of packets when MAC anti-spoofing is enabled. Currently, this command applies to the downstream PPPoE active discovery terminate (PADT) packet only. When a terminal goes offline, the PPPoE session heartbeat between the BRAS and the terminal is interrupted. In this case, the BRAS terminates the PPPoE session and sends a downstream PADT packet.

- By default, when receiving a downstream PADT packet, the access device deletes the dynamic MAC address binding entry of the corresponding PPPoE user. If the PPPoE user connects to another port of the access device, the user can immediately dial up from the new port and go online.
- When the **security anti-macspoofing ignore downstream-padt** command is configured, the access device does not process (that is, ignores) the downstream PADT packet. If the PPPoE user connects to another port of the access device, the user cannot immediately dial up from this new port, because the MAC address of the user terminal conflicts with the existing dynamic MAC address binding entry on the previous port. The user can dial up from a new port and go online only after the dynamic MAC address binding entry ages on the previous port.

Recovery of MAC Address Binding Entries

In V800R015C00, The contents of MAC address binding entries are stored as user data management (UDM) data. Recovery of MAC address binding entries is a feature by using which the system restores the UDM data to the memory when the system is restarted. Users do not need to dial up again to bind MAC addresses. Recovery of MAC address binding entries supports power-off recovery and non-power-off recovery. In non-power-off recovery, the memory space of the UDM data is not cleared. After the system is restarted, the system obtains the MAC address binding entries to implement MAC address binding. Non-power-off recovery requires no configuration and the working principle is simpler than that of the power-off recovery. The following describes the power-off recovery scenario.

- When the system is running properly, the UDM data is periodically compressed and backed up on the server through FTP/TFTP/SFTP. During automatic system backup, DHCP, DHCPv6, SLAAC, and PPPoE dialup users are forbidden to go online or offline. This is intended to avoid data conflicts. If a device power failure occurs during automatic system backup, the file stored on the server is incomplete. In such a case, the UDM data cannot be recovered after the system is restarted.
- When the system is restarted after a power failure, the system automatically downloads the backup data from the server and restores it after decompression. Because automatic download is performed during system startup, the upstream port may not be ready for automatic download and the download channel may not be available. In this case, automatic download cannot be smoothly carried out. The system makes attempts to download data from the server till the timeout time elapses. If no attempt is successful,

the system does not make any further attempts. During automatic download, data recovery, and data download attempts, dialup users are not allowed to go online or offline. This is to avoid data conflicts. Once automatic data backup is disabled during data download or data download attempts, users can go online and offline. If automatic data backup is disabled during data recovery after data download, users can go online and offline only after the UDM data is recovered.

- When the system configured with active/standby servers is restarted due to a power failure, the system will try to download data from the active server first. If the active server is not available, the system will try the standby server. When the file downloaded from the active server fails to be verified or is not the latest, the system will not download data from the standby server.
- The lease time of the recovered UDM data may be different from that of the original UDM data when the system time is changed in the following conditions: before a device power failure occurs without any automatic data backup; after the system is restarted due to a power failure while the UDM data has not been completely recovered.
- If a device power failure occurs after you run the **active configuration system** command but before the first UDM data backup is complete, MAC address binding entries cannot be correctly recovered after the system is restarted.

28.3.3 Configuring MAC Anti-spoofing

When users dynamically obtain IP addresses in PPPoE, DHCP, DHCPv6, or StateLess Address AutoConfiguration (SLAAC) dialup mode, the most frequently used MAC address security method is MAC anti-spoofing.

Prerequisites

Before the maximum number of MAC addresses that can be bound to a service stream is configured, the service port carrying this service stream must exist.

Context

The following uses an IPv4 network as an example to describe how to configure the MAC anti-spoofing feature. The configuration of this feature for IPv6 networks is similar to that for IPv4 networks. The user can configure this feature on IPv6 networks by referring to the example below.

Procedure

Run the **security anti-macspoofing enable** command to enable MAC anti-spoofing globally.

Step 1 Use one of the following methods to configure MAC anti-spoofing based on VLANs:

- Method 1: In global config mode, run the **security anti-macspoofing vlan** *vlanid* **enable** command to enable MAC anti-spoofing for a VLAN.
- Method 2: Use a VLAN service profile to configure MAC anti-spoofing.
  a. Run the **vlan service-profile** command to create a VLAN service profile and enter the VLAN service profile mode.
  b. Run the **security anti-macspoofing enable** command to enable MAC anti-spoofing for a VLAN.
  c. Run the **commit** command to make the VLAN service profile take effect.
  d. Run the **quit** command to exit the VLAN service profile mode.

e. Run the **vlan bind service-profile** command to bind the VLAN service profile created in step a to the VLAN.

NOTE

When multiple parameters to be configured for a large number of VLANs have the same value, configure these parameters in a VLAN service profile. Then bind the VLAN service profile to a desired VLAN. In this manner, the configuration workload for a single VLAN can be reduced. Method 2 applies to this scenario.

Step 2 Run the **security anti-macspoofing service-port***service-portid* **enable/disable** command to enable or disable MAC anti-spoofing based on service port.

Step 3 (Optional) Run the **security anti-macspoofing max-mac-count** command to configure the maximum number of MAC addresses that can be bound to a service port.

By default, a maximum number of eight MAC addresses can be bound to a service port. Perform this step when fewer than eight MAC addresses need to be bound.

Step 4 (Optional) Run the **security anti-macspoofing exclude** command to enable or disable source MAC address check for certain packets, such as Internet Group Management Protocol (IGMP) packets, in the case of MAC anti-spoofing.

This command applies only to IGMP packets. By default, the source MAC address of IGMP packets is not checked when MAC anti-spoofing is enabled.

- After the **undo security anti-macspoofing exclude IGMP** command is run, the access node checks the source MAC address carried in IGMP packets. The access node allows the IGMP packets to pass through only when the source MAC address is the same as the bound MAC address.
- After the **security anti-macspoofing exclude IGMP** command is run, the access node does not check the source MAC address carried in IGMP packets. The access node allows the IGMP packets to pass through regardless of whether the source MAC address is the same as the bound MAC address.

Step 5 (Optional) Enable power-off recovery of MAC address binding entries.

If you want users to go on line without dialup after a device power failure occurs, configure this function.

1. Run the **security user auto-backup enable** command to enable automatic data backup.
2. Run the **file-server auto-backup udm** command to configure the auto-backup server.
3. Run the **security user auto-backup period** command to configure the period for automatic data backup.
4. Run the **security user auto-load timeout** command to configure the timeout parameters for automatic data download. The timeout parameters include the total timeout time and the interval between each download attempt. If download is not finished before the timeout time elapses, the system stops data download.

Step 6 Query the configuration results. The following table lists the commands for querying MAC anti-spoofing configurations.

Queried Information	Command
Global MAC anti-spoofing configurations	**display security config**
MAC address filtering entries	**display security mac-filter**
Dynamically bound MAC addresses	**display security bind mac**

Queried Information	Command
Maximum number of MAC addresses that can be bound to a service port	**display security anti-macspoofing max-mac-count**
VLAN service profile configurations	**display vlan service-profile**

----End

Example

NOTE
In the following examples, information irrelevant with the configuration task is omitted in the output of the **display security config** command. For complete command output, see *Command Reference*.

An FTTH user accesses the Internet in PPPoE dialup mode. Assume that the service port index is 1 and the S-VLAN ID is 1000. To enable MAC anti-spoofing for the user, and use the default value for the maximum number of MAC addresses that can be bound to the service port, do as follows:

```
huawei(config)#security anti-macspoofing enable
huawei(config)#security anti-macspoofing vlan 1000 enable
huawei(config)#display security config
  Anti-macspoofing function                    : enable
  Anti-macspoofing control-protocol IPv6oE function : disable
  Packet unaffected by anti-macspoofing          : IGMP
```

To disable MAC anti-spoofing of service port 1 in a trusted network, do as follows:

```
huawei(config)#security anti-macspoofing service-port 1 disable
```

An FTTH user accesses the Internet in DHCPv6 dialup mode. Assume that the service port index is 2 and the S-VLAN ID is 1000. To enable MAC anti-spoofing for the user, and set the maximum number of MAC addresses that can be bound to the service port to 3, do as follows:

```
huawei(config)#security anti-macspoofing enable
huawei(config)#security anti-macspoofing vlan 1000 enable
huawei(config)#security anti-macspoofing control-protocol ipv6oe
huawei(config)#security anti-macspoofing max-mac-count service-port 2 3
huawei(config)#display security config
  Anti-macspoofing function                    : enable
  Anti-macspoofing control-protocol IPv6oE function : enable
  Packet unaffected by anti-macspoofing          : IGMP
```

An FTTH user accesses a network in DHCP mode to obtain Internet access, voice, and video services. Assume that the service port indexes for the three types of services are 3, 4, and 5, and the S-VLAN IDs are 1000, 1001, and 1002. To enable MAC anti-spoofing for the user, and use the following settings for the user:

- Set the maximum number of MAC addresses that can be bound to each service port to 2.
- Use multicast mode for video services.
- Use the default setting for source MAC address check, that is, the source MAC address of IGMP packets is not checked in the case of MAC anti-spoofing.

do as follows:

```
huawei(config)#security anti-macspoofing enable
huawei(config)#vlan service-profile profile-id 1
huawei(config-vlan-srvprof-1)#security anti-macspoofing enable
huawei(config-vlan-srvprof-1)#commit
huawei(config-vlan-srvprof-1)#quit
huawei(config)#vlan bind service-profile 1000-1002 profile-id 1
huawei(config)#security anti-macspoofing max-mac-count service-port 3 2
huawei(config)#security anti-macspoofing max-mac-count service-port 4 2
huawei(config)#security anti-macspoofing max-mac-count service-port 5 2
huawei(config)#display security config
  Anti-macspoofing function                    : enable
  Anti-macspoofing control-protocol IPv6oE function : disable
  Packet unaffected by anti-macspoofing          : IGMP
```

28.3.4 Maintenance and Diagnosis

To facilitate location of malicious users, the access node provides statistics collection, logging, and event reporting functions for MAC address conflicting packets and MAC spoofing packets. After receiving such packets, the access node reports appropriate events.

The following methods can be used to locate malicious users when MAC anti-spoofing is enabled on the access node:

- Run the **display security conflict statistic** command to collect statistics on MAC spoofing packets among user data packets. The number of MAC spoofing packets is indicated by **MAC conflict packets number** in the command output.
- Run the **display security conflict log** command to record MAC address conflicting packets and MAC spoofing packets among user protocol packets.
- After detecting MAC address conflicting packets or MAC spoofing packets, the access node reports events to facilitate location of malicious users. Table 28-6 lists MAC anti-spoofing events that are reported by the access node.

Table 28-6 MAC anti-spoofing events

Type	Event
Events related to protocol packets	• 0x28000033 The user of this port uses the MAC address that is already bound to another user or uses the MAC address that is not bound to this user • 0x28000034 The user of this PON port uses the MAC address that is already bound to another user or uses the MAC address that is not bound to this user • 0x28000024 The invalid ARP packet with the source MAC address different from the SHA field is received from the user on the port • Distributed mode: 0x28000025 The invalid ARP packet with the source MAC address different from the SHA field is received on the GPON port from the user • Profile mode: 0x28000026 The invalid ARP packet with the source MAC address different from the SHA field is received from the user on the GPON port
Events related	• 0x28000021 The data packet with the source MAC address different

Type	Event
to data packets	from the MAC address bound to this user is received on the port • Distributed mode: 0x28000022 The data packet with the source MAC address different from the MAC address bound to this user is received on the GPON port • Profile mode: 0x28000023 The data packet with the source MAC address different from the MAC address bound to this user is received on the GPON port

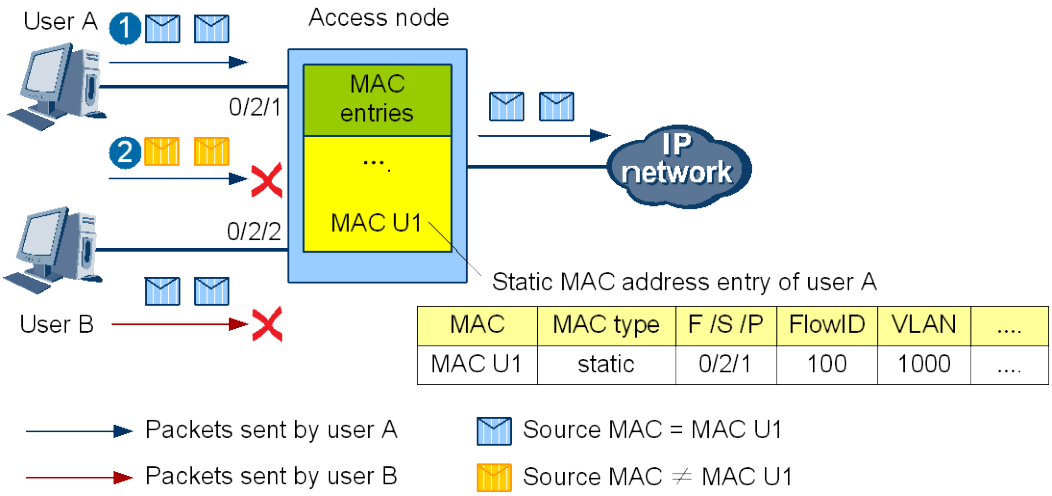
28.4 Static MAC Address Binding

When users' IP addresses are manually configured, the most frequently used method for anti-user MAC address spoofing is static MAC address binding.

28.4.1 Principle

After static MAC addresses are bound to a service stream, a user's MAC address can be protected against forgery and the user can be prevented from forging the MAC addresses of other users or upper-layer devices. Figure 28-6 shows the principle of static MAC address binding.

Figure 28-6 Static MAC address binding



The port of user A is 0/2/1 and the service stream index (FlowID) is 100. A static MAC address entry (MAC U1) is configured for user A on the access node, and the number of dynamic MAC addresses that can be learned by the service stream of user A is set to 0.

1. The packets sent by user A using MAC U1 as the source MAC address can pass through the access node.
2. The packets sent by user A using other source MAC addresses are discarded by the access node. In this manner, user A cannot forge the source MAC addresses of other users or upper-layer devices. The principle is that the service stream of user A cannot

learn any dynamic MAC address after the number of learnable dynamic MAC addresses is set to 0. Therefore, the access node can forward packets only based on the configured static MAC address. If the packets sent by user A carry a source MAC address other than the configured static MAC address (for example, the MAC address of an upper-layer device), the packets will be discarded by the access node.

3. The packets sent by user B using MAC U1 as the source MAC address are discarded by the access node, because the source MAC address is the same as the static MAC address configured for user A. In this manner, user A's source MAC address can be protected against forgery. The principle is that the static MAC address has a higher priority than the dynamic MAC address. If a MAC address has been statically configured for a port, this MAC address will not be learned by other ports as a dynamic MAC address in the same VLAN.

28.4.2 Configuring Static MAC Address Binding

When users' IP addresses are manually configured, the most frequently used method for anti-user MAC address spoofing is static MAC address binding.

Prerequisites

Before static MAC addresses and the maximum number of learnable dynamic MAC addresses are configured for a service stream, the service port carrying this service stream must exist.

Procedure

Run the **mac-address static** command to add a static MAC address.

Step 1 Run the **mac-address max-mac-count** command to set the maximum number of dynamic MAC addresses that can be learned by a service stream to 0.

Step 2 Query the configuration results. The following table lists the commands for querying the configurations of static MAC address binding.

Queried Information	Command
MAC address table of the device, including all static MAC addresses and the source MAC addresses of service streams	**display mac-address**
Maximum number of dynamic MAC addresses that can be learned by a service stream	**display mac-address max-mac-count**

----End

Example

Assume that the index of a service stream of an enterprise private line user is 100 and the source MAC address is 1010-1010-1010. To bind the static MAC address 1010-1010-1010 to the service stream, do as follows:

```
huawei(config)#mac-address static service-port 100 1010-1010-1010
huawei(config)#mac-address max-mac-count service-port 100 0
```

```
huawei(config)#display mac-address service-port 100
  ------------------------------------------------------------------------
   SRV-P BUNDLE TYPE MAC          MAC TYPE F /S /P   VPI  VCI   VLAN ID
   INDEX INDEX
  ------------------------------------------------------------------------
     100    -  vdl  1010-1010-1010 static   0/ 2/ 0   -    -       1000
  ------------------------------------------------------------------------
  Total: 1
  Note: F--Frame, S--Slot, P--Port, F/S/P indicates PW Index for PW,
       A--The MAC address is learned or configured on the aggregation port,
       VPI indicates GEM Port ID for GPON,
       v/e--vlan/encap, pri-tag--priority-tagged,
       ppp--pppoe, ip--ipoe, ip4--ipv4oe, ip6--ipv6oe

huawei(config)#display mac-address max-mac-count service-port 100
  ------------------------------------------------------------------------
   SRV-P TYPE  F /S /P  VPI  VCI   VLAN ID FLOWTYPE   FLOWPARA   LEARNABLE
   INDEX                                              MAC NUMBER
  ------------------------------------------------------------------------
     100 vdl   0/ 2/ 0  -    -       1000 -         -               0
  ------------------------------------------------------------------------
  Total: 1
  Note: F--Frame, S--Slot, P--Port,
       A--The MAC address is learned or configured on the aggregation port,
       VPI indicates GEM Port ID for GPON,
       v/e--vlan/encap, pri-tag--priority-tagged,
       ppp--pppoe, ip--ipoe, ip4--ipv4oe, ip6--ipv6oe
```

28.5 Static MAC Address Filtering

To prevent users from forging the MAC addresses of upper-layer devices or forging certain well-known MAC addresses, set the well-known MAC addresses and the MAC addresses of the upper-layer devices to the MAC addresses to be filtered. The static MAC address filtering feature includes static source MAC address filtering and static destination MAC address filtering.

28.5.1 Principle

Static Source MAC Address Filtering

The MAC addresses of upper-layer devices are added to source MAC address filtering entries. Then the MAC addresses of the upper-layer devices cannot be used by malicious users as the source MAC addresses to send packets. When users' IP addresses are manually statically configured, static source MAC address filtering is the most frequently used method for preventing upper-layer device MAC address spoofing.

The basic principle for source MAC address filtering is that packets are discarded if they carry source MAC addresses that are the same as those in source MAC address filtering entries. To achieve static source MAC address filtering, the MAC addresses of upper-layer devices must be manually added to source MAC address filtering entries.

Upper-layer device MAC address spoofing can also be prevented by configuring static MAC addresses for upstream ports. Compared with the method of preventing upper-layer device MAC address spoofing by configuring static MAC addresses for upstream ports, static source MAC address filtering (including dynamic and static) has the following advantages and disadvantages:

- The static source MAC address filtering feature cannot be configured based on service streams or ports, which has advantages as well as disadvantages.
  - Advantage: The mapping between the upper-layer device and upstream port does not need to be known before configuration, which reduces the configuration workload. After the configuration is complete, the mapping between the upper-layer device and upstream port can be changed.
  - Disadvantage: The static source MAC address filtering feature can only prevent the MAC addresses of upper-layer devices from being used by malicious users. The upper-layer device connected to an upstream port, however, can use the MAC address of the upper-layer device connected to another upstream port as the source MAC address to send packets. Generally, upstream ports are trustable, so this disadvantage can be ignored.
- Only a small number of static source MAC addresses can be filtered by the access node. In normal cases, there are only a few upper-layer devices, so the number of MAC addresses of upper-layer devices is also small. Therefore, this disadvantage can also be ignored.

Based on the comparison, static source MAC address filtering is a better method for preventing upper-layer device MAC address spoofing than the method of configuring static MAC addresses for upstream ports.

Static Destination MAC Address Filtering

The MAC addresses of upper-layer devices are added to destination MAC address filtering entries. Then the MAC addresses of the upper-layer devices cannot be used by malicious users as the destination MAC addresses in packets, so that the malicious users cannot access the upper-layer device by sending unicast packets.

28.5.2 Configuring Static MAC Address Filtering

To prevent users from forging the MAC addresses of upper-layer devices or forging certain well-known MAC addresses, set the well-known MAC addresses and the MAC addresses of the upper-layer devices to the MAC addresses to be filtered. The static MAC address filtering feature includes static source MAC address filtering and static destination MAC address filtering.

Procedure

- Configure static source MAC address filtering.

  Run the **security mac-filter source** command to configure the source MAC address to be filtered out.
- Configure static destination MAC address filtering.

  Run the **security mac-filter destination** command to configure the destination MAC address to be filtered out.
- Run the **display security mac-filter** command to query MAC address filtering entries.

----End

Example

Assume that the MAC address of an upper-layer device (for example, the BRAS) is 1010-1010-3020. To add this MAC address to the source MAC address filtering table of the access node, do as follows:

```
huawei(config)#security mac-filter source 1010-1010-3020
huawei(config)#display security mac-filter
{ <cr>|destination<K>|dynamic<K>|source<K> }:

 Command:
        display security mac-filter
 ------------------------------------------------------------------------------
 Index      MAC-Address          Type          Filter-Mode      VLAN
 ------------------------------------------------------------------------------
   0       1010-1010-3020       static        source             -
 ------------------------------------------------------------------------------
 Total: 1
```

Assume that the MAC address of a well-known website is 1010-1010-4020. To add this MAC address to the destination MAC address filtering table of the access node, do as follows:

```
huawei(config)#security mac-filter destination 1010-1010-4020
huawei(config)#display security mac-filter
{ <cr>|destination<K>|dynamic<K>|source<K> }:

 Command:
        display security mac-filter
 ------------------------------------------------------------------------------
 Index      MAC-Address          Type          Filter-Mode      VLAN
 ------------------------------------------------------------------------------
   0       1010-1010-3020       static        source             -
   1       1010-1010-4020       static        destination         -
 ------------------------------------------------------------------------------
 Total: 2
```

28.6 MAC Anti-Duplicate

Before the source MAC address of a port is aged, the access node learns the source MAC address of the port from another port and then updates the mapping between the source MAC address and the port in the MAC address table. This process is called MAC address duplicate (also known as MAC address flapping), because it can be regarded as if the access node copied the MAC address from one port to another port (the MAC address flaps from one port to another port). To prevent the MAC addresses of authorized users or upper-layer devices from being duplicated by malicious users, the access node supports the MAC anti-duplicate function.

28.6.1 Introduction

MAC address duplicate is related to MAC address learning. The access node receives packets through port A and then learns a source MAC address from the packets. If the learned source MAC address maps port B in the MAC address table, the access node will change the port

that maps the learned source MAC address from port B to port A. During this process, it can be considered that the access node duplicates the MAC address from port B to port A. Therefore, this process is called MAC address duplicate.

MAC address duplicate has four types:

- Type 1: The MAC address learned by a user-side port is duplicated to another user-side port.
- Type 2: The MAC address learned by a user-side port is duplicated to a network-side port.
- Type 3: The MAC address learned by a network-side port is duplicated to a user-side port.
- Type 4: The MAC address learned by a network-side port is duplicated to another network-side port.

NOTE

User-side ports refer to user ports and cascading ports, and network-side ports refer to upstream ports.

The access node supports the MAC anti-duplicate function to prevent the MAC addresses of authorized users or upper-layer devices from being duplicated by malicious users.

The principle of the MAC anti-duplicate function is as follows: The access node receives packets through port A, learns the source MAC address, and looks up the source MAC address in the MAC address table. If the source MAC address maps port B in the MAC address table, the access node enabled with MAC anti-duplicate will determine whether to duplicate the source MAC address from port B to port A based on types of control boards, service boards and ports. If duplicate is forbidden, the access node will discard the packets before the source MAC address is aged.

28.6.2 Principle

The MAC anti-duplicate feature can be implemented in two ways: through software or through hardware. The feature is configured using the **security anti-macduplicate** command in both implementation modes.

Software MAC Anti-Duplicate

When the device implements MAC anti-duplicate through software, the MAC address duplicate types permitted and forbidden by the device are as shown in the following table.

MAC Address Duplicate Type	security anti-macduplicate enable	security anti-macduplicate disable
Type 1	Permit	Permit
Type 2	Permit	Permit
Type 3	Forbid	Permit
Type 4	Forbid	Permit

Hardware MAC Anti-Duplicate

Hardware MAC anti-duplicate is related to the MAC address learning priority of the port. The following tables show the details.

Table 28-7 MAC address learning priority of user-side ports lower than that of network-side ports

MAC Address Duplicate Type	security anti-macduplicate enable	security anti-macduplicate disable
Type 1	Forbid	Permit
Type 2	Permit	Permit
Type 3	Forbid	Forbid
Type 4	Permit	Permit

Table 28-8 MAC address learning priority of user-side ports higher than that of network-side ports

MAC Address Duplicate Type	security anti-macduplicate enable	security anti-macduplicate disable
Type 1	Forbid	Permit
Type 2	Forbid	Forbid
Type 3	Permit	Permit
Type 4	Permit	Permit

Table 28-9 MAC address learning priority of user-side ports the same as that of network-side ports

MAC Address Duplicate Type	security anti-macduplicate enable	security anti-macduplicate disable
Type 1	Forbid	Permit
Type 2	Forbid	Permit
Type 3	Forbid	Permit
Type 4	Forbid	Permit

28.6.3 Configuring MAC Anti-Duplicate

Before the source MAC address of a port is aged, the access node learns the source MAC address of the port from another port and then updates the mapping between the source MAC

address and the port in the MAC address table. This process is called MAC address duplicate (also known as MAC address flapping), because it can be regarded as if the access node copied the MAC address from one port to another port (the MAC address flaps from one port to another port). To prevent the MAC addresses of authorized users or upper-layer devices from being duplicated by malicious users, the access node supports the MAC anti-duplicate function.

Procedure

Run the **security anti-macduplicate** command to enable the MAC anti-duplicate function.

Step 1 Run the **display security config** command to query the configuration results.

----End

Example

NOTE

In the following example, information irrelevant with the configuration task is omitted in the output of the **display security config** command. For complete command output, see *Command Reference*.

To enable the MAC anti-duplicate function for the access node, do as follows:

```
huawei(config)#security anti-macduplicate enable
huawei(config)#display security config
  Anti-macduplicate function                  : enable
```

28.6.4 Maintenance and Diagnosis

When certain control boards work with service boards , the system generates the 0x64000001 Source MAC got by non-network-side ports is the same as one learnt by network-side ports event if the source MAC address of the packet received on the user port or cascade port is the same as the MAC address learned on the upstream port. For details, see Specification.

28.7 VMAC

A Virtual MAC address is a network-wide unique MAC address generated by the access device according to certain rules. VMAC can be used to prevent untrusted MAC addresses from accessing carriers' networks and prevents malicious users from forging the source MAC addresses of trusted users or upper-layer devices.

28.7.1 Introduction

Context

Each MAC address on a Layer 2 network must be unique. The MAC address allocation mechanism ensures global uniqueness of each address. However, hackers use scanning tools to obtain existing MAC addresses, which allow hackers to impersonate genuine users. The impersonation of a MAC address is known as MAC spoofing. Duplicate MAC addresses exist in MAC spoofing; the same MAC address appears on different ports of a switch, causing a MAC address transfer on the switch. As a result, data is sent to the hacker's device instead of to the genuine user. There are two types of MAC spoofing:

- MAC spoofing to upstream service servers (such as a BRAS, DHCP server, or trunk gateway)
- MAC spoofing to downstream users.

Generally, operators control the aggregation network directly, which protects against MAC spoofing or duplication. The end-user system, constituted by a large number of users, is hard to control, because the MAC addresses of end-users are not trustworthy to carriers. Virtual media access control (VMAC) provides carriers another way to protect against MAC spoofing and duplication.

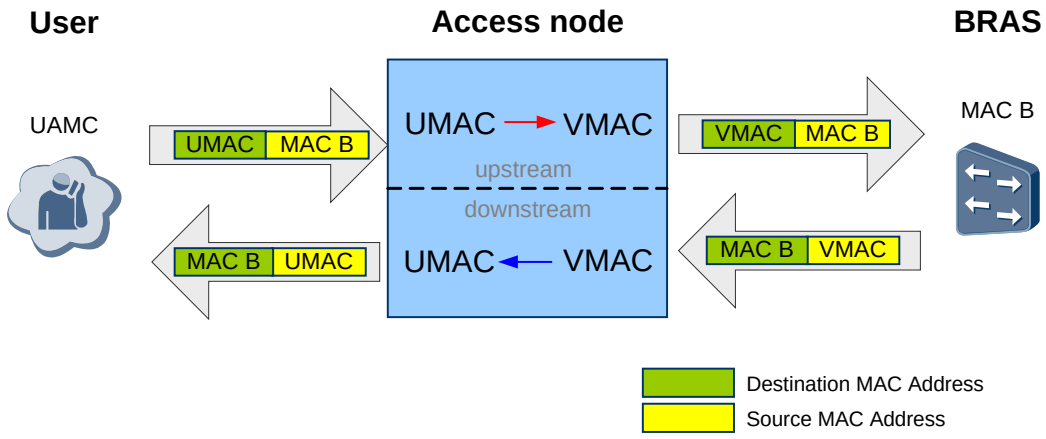
Definition

A VMAC address is a network-wide unique MAC address generated by the access device according to certain rules. Because the VMAC address is generated by the access device (the access node), it is considered trustworthy by carriers.

After the VMAC function is enabled, the access device, in the upstream direction, replaces untrustworthy source MAC addresses in packets received from the user side with trusted VMAC addresses, and then forwards these packets to the upstream network. In the downstream direction, the access device restores the VMAC addresses in packets received from the network side to the actual MAC addresses of users, and then sends these packets to user ports.

The following figure shows the principle of VMAC.

Figure 28-7 Principle of VMAC



The access device supports 1:1 VMAC and N:1 VMAC.

- 1:1 VMAC: The access device converts each user MAC address (UMAC) into a unique VMAC address. UMAC and VMAC are in a 1:1 mapping.
- N:1 VMAC: The access device converts a group of UMACs into a unique VMAC address. UMACs and VMAC are in an N:1 mapping.

Benefits

Benefits for Carriers

- The VMAC function improves network security. VMAC ensures network-wide uniqueness of MAC addresses and prevents issues that may arise from MAC address

conflicts. In addition, VMAC prevents untrustworthy user MAC addresses from entering a carrier's network and protects the carrier network from MAC spoofing.

- 1:1 VMAC identifies users by providing precise information about a user line.
- N:1 VMAC allows the access node to accommodate more users by using the same MAC address space.

Benefits for Users

The VMAC function protects users from MAC address conflicts and MAC spoofing.

28.7.2 1:1 VMAC Principles

In 1:1 VMAC, the access device converts each user MAC address (UMAC) into a unique VMAC address. UMAC and VMAC are in a 1:1 mapping.

VMAC Address Allocation Mode

- For Ethernet packets, such as PPPoE, IPoE, and ARP packets, the access device generates VMAC addresses for replacing UMAC addresses according to the VMAC address format. When VMAC is enabled, the access device automatically allocates VMAC address pool to all registered service boards, starting from the service board with the smallest slot ID. Each board is allocated with the number of VMAC addresses corresponding to the number of ports and each port is allocated with the number of VMAC addresses according to the specification. Each time a new user goes online under a port, the access device performs MAC address translation (MAT) on the UMAC, translating the UMAC into a VMAC that has been allocated to the port.
- For PPPoA packets, the access device allocates the VMAC out of the MAC address pool that has been configured for users, and translates the UMAC into this VMAC. Users can configure the MAC address pool but cannot specify MAC address segments for specific ports or service boards.

Format of 1:1 VMAC

The following table describes the general format of 1:1 VMAC.

Table 28-10 General format of 1:1 VMAC

MAC Address Field	Description
Bits 47-42	Reserved bits, configurable by users through commands for special designation purposes.
Bit 41	Fixedly set to 1, indicating a local MAC address.
Bit 40	Fixedly set to 0, indicating a unicast MAC address.
Bits 39-21	DSLAM ID configured by the user, identifying the access device in a network. This field must be set to a unique value to ensure that different access devices will generate unique VMAC addresses, even if the other fields of their MAC addresses are not unique.
Bits 20-15	Slot ID, indicating the ID of the slot where the user locates.
Bits 14-6	Port ID, indicating the ID of the port to which the user belongs.

MAC Address Field	Description
Bits 5-0	The index allocated to online users by the access device.

The format of 1:1 VMAC in GPON access is different from the general format of 1:1 VMAC. The differences are the length of some fields, and the addition of the ONT ID field for identifying different ONTs connected to the same GPON port. The following table describes the format of 1:1 VMAC in GPON access.

Table 28-11 Format of 1:1 VMAC in GPON access

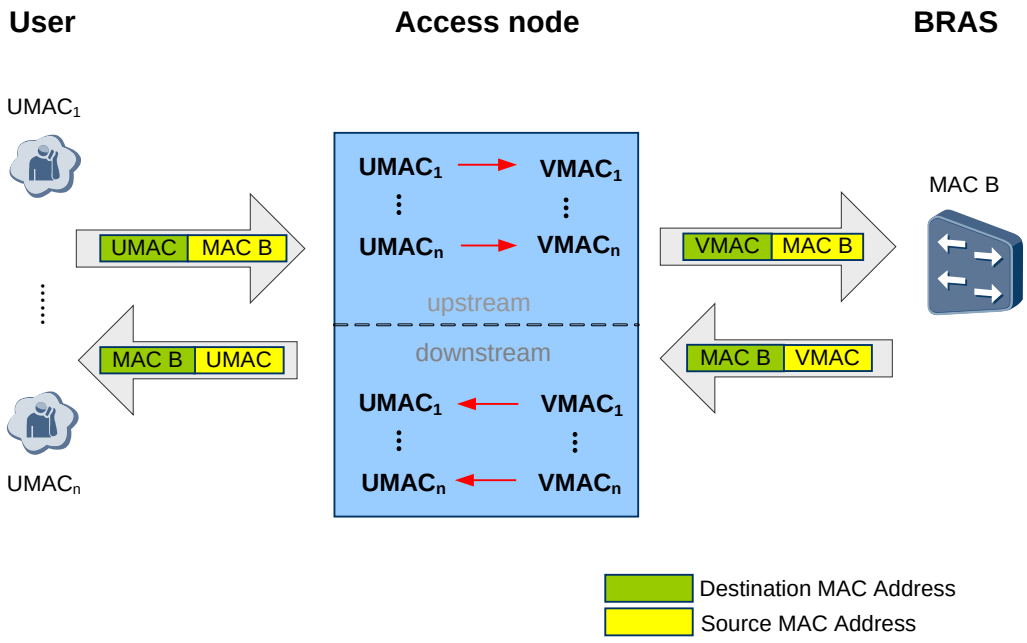
MAC Address Field	Description
Bits 47-42	Reserved bits, configurable by users through commands for special designation purposes.
Bit 41	Fixedly set to 1, indicating a local MAC address.
Bit 40	Fixedly set to 0, indicating a unicast MAC address.
Bits 39-24	OLT ID configured by the user, identifying the access device in a network. This field must be set to a unique value to ensure that different access devices will generate unique VMAC addresses, even if the other fields of their MAC addresses are not unique.
Bits 23-18	GPON board slot ID, indicating the ID of the slot where the user locates. **NOTE** The displayed value is based on the configuration of the **vmac slot-numbering** command. • **logical**: The ID of the slot to which users belong is equal to the ID of the actual slot plus one. • **physical**: The ID of the slot to which users belong is equal to the ID of the actual slot.
Bits 17-13	GPON port ID, indicating the ID of the GPON port to which the user belongs. **NOTE** The displayed value is based on the configuration of the **vmac port begin** command. • **0**: The ID of the port to which users belong is equal to the ID of the actual port. • **1**: The ID of the port to which users belong is equal to the ID of the actual port plus one.
Bits 12-3	ONT ID, identifying the ONT to which the user is connected. **NOTE** The value of this field is the actual ONT ID plus 1.
Bits 2-0	The index allocated to online users by the access device.

Principle of 1:1 VMAC for PPPoE/IPoE

NOTE

1:1 VMAC works in the same way with PPPoE, IPoE, ARP, ND, DHCP, ETH OAM LTM, and ETH OAM LTR packets. The following section uses PPPoE/IPoE packets as an example to describe the principle of 1:1 VMAC.

Figure 28-8 Processing flow of 1:1 VMAC for PPPoE/IPoE



- **In the upstream direction:**

  The access device processes the packet differently after receiving the packet from the user.

  - If the packet carries a new UMAC as the source MAC address, the access device translates the UMAC into a VMAC generated by the system and sends the packet to the upstream network. At the same time, the access device adds a corresponding entry to the system's UMAC-VMAC mapping table.
  - If the access device finds that the source MAC address of the packet already exists in the UMAC-VMAC mapping table (the system has allocated a VMAC for this source MAC address), the system updates the aging flag for this MAC address, translates the MAC address into the VMAC, and sends the packet to the upstream network.
  - If the system finds that the number of online users has reached the maximum number of VMAC addresses supported, for example, online users on a port have reached the maximum number of VMAC addresses supported by the port, the system drops the received packet from the new subscriber on the corresponding service board.

- **In the downstream direction:**

  The access device looks up the address resolution list (ARL) by using the VLAN+destination MAC address (DMAC) entry, and obtains information about the port for forwarding the packet. Here, the DMAC address is the VMAC address. The system then looks up the VMAC table to obtain the VMAC-UMAC mapping, and translates the VMAC address into the UMAC address when forwarding the packet to the user port.

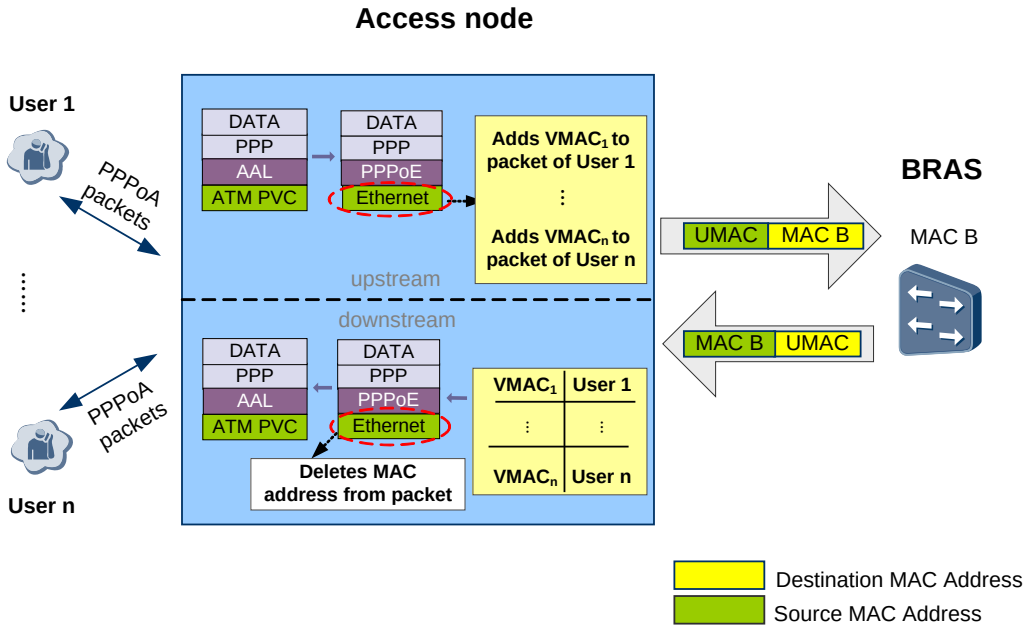
NOTE

- For unicast packets, the system performs MAT on the MAC addresses in upstream and downstream directions for trusted users.
- For multicast packets: The VMAC function does not take effect. The access device does not perform MAT on the MAC addresses on the multicast MAC addresses.
- For broadcast packets: The system performs MAT on the MAC addresses only in the upstream direction for trusted users, that is, the system translates UMAC into VMAC. The system broadcasts the packets in the downstream direction.
- The user source MAC address exists not only in the Ethernet header but also in the data field. The access device translates the MAC address in both the Ethernet header and the data field.

Principle of 1:1 VMAC for PPPoA

PPPoA packets differ from PPPoE/IPoE packets in that PPPoA packets do not carry MAC addresses. The access device supports 1:1 VMAC for PPPoA packets by translating PPPoA packets into PPPoE packets. A VMAC-enabled access device adds VMAC addresses to the received PPPoA packets. The following figure shows the principle of 1:1 VMAC for PPPoA.

Figure 28-9 Processing flow of 1:1 VMAC for PPPoA



- **In the upstream direction:**

  The access device converts the received PPPoA packet into a PPPoE packet, and fills in an allocated VMAC address as the source MAC address of each PPPoA session (identified by a unique session ID). The system then forwards the packet to the upstream network.

- **In the downstream direction:**

  The DMAC address of each PPPoA session is the VMAC address. The access device looks up the ARL by using the VLAN+VMAC entry to obtain information about the port for forwarding the packet, and deletes the MAC address from the packet when forwarding the packet to the port.

Mechanism for Coping with Changes of UMAC-VMAC Mappings

The access device saves the established UMAC-VMAC mappings into a UMAC-VMAC mapping table. The system does not save a UMAC-VMAC entry in the following conditions:

- The system is powered off, or the system resets after it is powered off.
- The aging time of a UMAC-VMAC entry expires.
- The DHCP user goes offline initiatively.
- The PPPoE user is disconnected or goes offline.

When the user goes online again, the system will the system will generate a new UMAC-VMAC entry, and the VMAC address allocated to the user may differ from the user's previous VMAC address.

NOTE

When the system or a service board is reset, or a port is activated:

- The system saves the UMAC-VMAC entry for the DHCP user. For the same DHCP user, the system uses the same VMAC address for performing MAT after a system reset, board reset, or port activation.
- The system does not save the UMAC-VMAC entry for the PPPoE user. The same PPPoE user may have a different VMAC after a system reset, board reset, or port activation.

VMAC Aging Mode

The system releases VMAC addresses that have not been used for a certain period of time. Released VMAC addresses can be allocated to other users.

The system supports two VMAC aging modes.

- **MAC learning mode**

  The aging time can be set by using the **mac-address timer** command. The system periodically checks for packets. If the system does not detect any packet, whether sent or received, carrying the VMAC address within twice the configured aging time, the system automatically releases the VMAC address.

- **DHCP mode**

  In this mode, VMAC address of a user will not age if the IP address of a user is allocated using DHCP. The VMAC ages only when the user's IP address is released or is not renewed after the lease expires.

  The DHCP aging mode applies only to the DHCP dialup service with the **multi-mac** MAC address allocation mode. This aging mode avoids frequent changes of system entries because it maintains the mapping between the DHCP user's IP address and VMAC address before the IP address is released. For example, after a computer wakes up from hibernation, the computer will not perform DHCP dialup again. At this moment, the IP address allocated to the computer through DHCP is not released, so the computer can still use the mapping VMAC address.

28.7.3 N:1 VMAC Principles

The access device translates a group of user MAC (UMAC) addresses into a unique VMAC address. UMAC addresses and the VMAC address are in an N:1 mapping.

VMAC Address Allocation Mode

- When the access device is not configured with an xPON protection group, the access device allocates a MAC address, in the N:1 VMAC address format, to each service board. The users connected to the same board will share this MAC address as the VMAC address for MAC address translation (MAT). Table 28-12 shows the format of an N:1 VMAC address.

Table 28-12 Format of an N:1 VMAC address

Slot ID	Formula	Illustration
Service board slot ID < control board slot ID	Board MAC address = System bridge MAC address + 9 + Service board slot ID	Assume that: • The slot ID of main control board is 0/9. • System bridge MAC address is 0x0000-0000-0001. The board MAC addresses get based on the formula for example are as follows. • The board MAC address of slot 0/1: 0x0000-0000-0001+9+1=0x0000-0000-000B • The board MAC address of slot 0/11: 0x0000-0000-0001+9+11-2=0x0000-0000-0013
Service board slot ID > control board slot ID	Board MAC address = System bridge MAC address + 9 + Service board slot ID - 2	

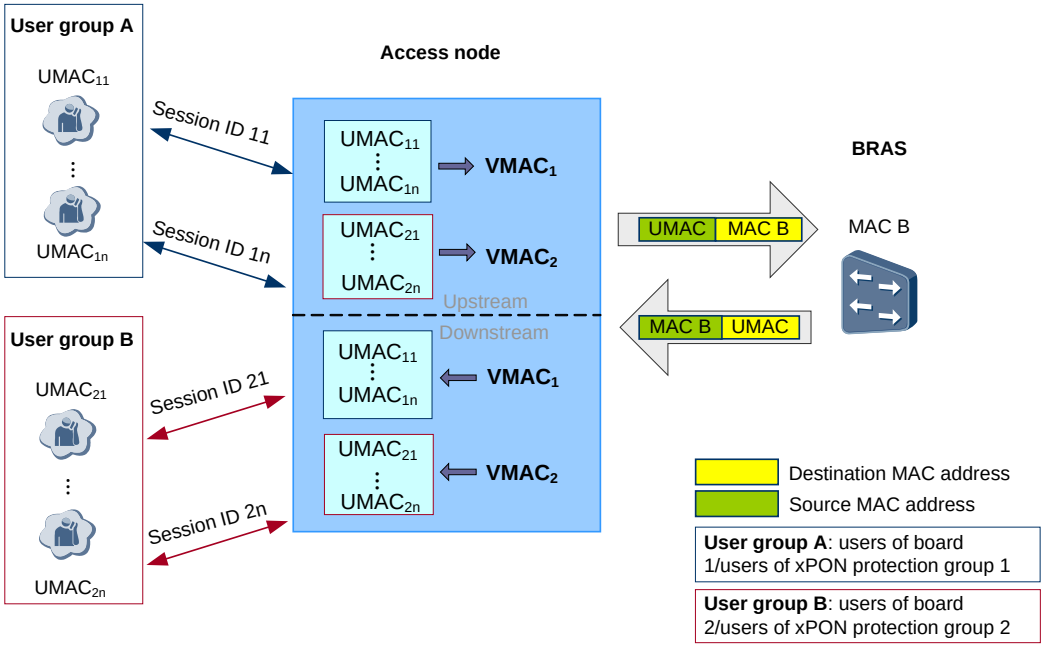
- When the access device is configured with an xPON protection group, the access device cannot perform MAT by directly using the MAC address of a service board if the protection group includes two service boards. The VMAC address used by the access device in this case for MAT is selected from a separately configured MAC address pool. The access device issues the VMAC address to the service board to which the user is connected and performs MAT on the service board.

NOTE

The address allocation mode of N:1 VMAC differs from that of 1:1 VMAC in that the MAC address allocated to each service board in N:1 VMAC will not be aged.

Principle of N:1 VMAC for PPPoE

Figure 28-10 Processing flow of N:1 VMAC for PPPoE



- **In the upstream direction:**

  For user packets that are received from the same service board (or the same xPON protection group), the access device translates the UMAC addresses of these packets into the VMAC address allocated to the service board, and forwards the packets to the upstream network.

- **In the downstream direction:**

  The destination MAC addresses of packets sent to the PPPoE sessions of the same service board are the VMAC address. The access device forwards the packets to the service board according to the mapping entry in the address resolution list (ARL). The service board forwards the packets to users according to the PPPoE session ID; at the same time, the access device translates the VMAC address in the packets into the respective UMAC addresses.

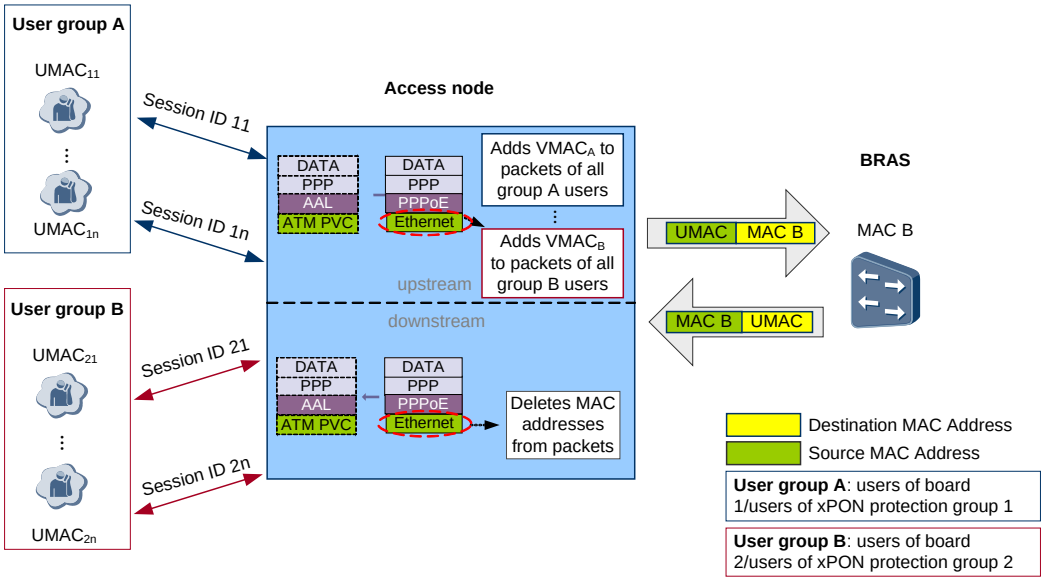
NOTE

- After N:1 VMAC is enabled, the system performs MAT only on the unicast packets of trusted users. The system does not perform MAT on the multicast or broadcast packets of these users.
- The user source MAC address exists not only in the Ethernet header but also in the data field. The system translates the source MAC address in both the Ethernet header and the data field.

Principle of N:1 VMAC for PPPoA

PPPoA packets differ from PPPoE/IPoE packets in that PPPoA packets do not carry MAC addresses. An VMAC-enabled access device adds VMAC addresses to the received PPPoA packets. The following figure shows the principle of N:1 VMAC for PPPoA.

Figure 28-11 Processing flow of N:1 VMAC for PPPoA



- **In the upstream direction:**

  For user packets that are received from the same service board (or the same xPON protection group), the access device fills in the source MAC addresses of these packets with the VMAC address allocated to the service board, and forwards the packets to the upstream network.

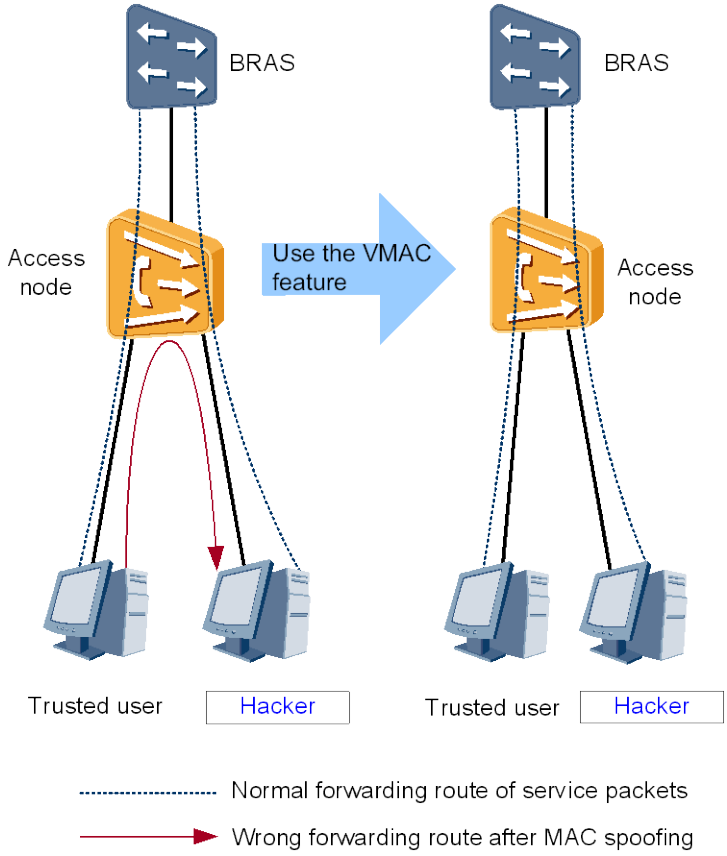
- **In the downstream direction:**

  The destination MAC addresses of packets sent to the PPPoA sessions of the same service board are the VMAC address. The access device deletes the MAC addresses from the packets, and forwards the packets to users according to the session ID.

28.7.4 Application

This topic provides an example of using the VMAC feature on a broadband remote access server (BRAS) to describe how the VMAC feature resolves MAC address spoofing and duplication.

Figure 28-12 Using the VMAC feature on a BRAS to resolve MAC spoofing



- The left part of the figure shows the situation before the VMAC feature is enabled. If the hacker forges the MAC address of the BRAS and sends a packet carrying this MAC address as the source MAC address, the access device saves the mapping between this MAC address and the hacker's port as an entry in the MAC address table. As such, the access device may forward the hacker a trusted user's packet that is destined for the BRAS.
- The right part of the figure shows the situation after the VMAC feature is enabled. Even if the hacker spoofs the MAC address of the BRAS, the hacker will not receive the packet of the trusted user. This is because the access device has replaced the source MAC address in the packet with a VMAC address, and the VMAC address differs from the actual MAC address of any device in the network.

The VMAC feature resolves MAC address duplication in a similar manner. The access device either translates the source MAC addresses of users into unique VMAC addresses (1:1 VMAC) one by one, according to the sequence in which the users go online, or translates the source MAC addresses into the VMAC addresses allocated to the board to which the users are connected (N:1 VMAC). The access device then forwards the packets to the upstream network, regardless of whether the MAC addresses repeat among the users.

28.7.5 Configuring 1:1 VMAC

In 1:1 VMAC, the system generates trusted virtual MAC addresses (VMACs) according to specified rules to replace source MAC addresses of end-users. 1:1 VMAC prevents users that have untrustworthy MAC addresses from accessing carriers' networks and is an effective

countermeasure to MAC spoofing. In addition, a carrier can directly locate a user in the carrier's network by the VMAC and obtain precise user line information.

Prerequisites

The VMAC function of all packets, except PPPoA packets, conflicts with anti-MAC spoofing. When VMAC is enabled for all packets (except PPPoA), make sure that anti-MAC spoofing is not enabled at the same time. You can run the **display security config** command to query the status of anti-MAC spoofing.

Context

Table 28-13 System defaults of 1:1 VMAC parameters

Parameter	Default Value
VMAC state	disable
VLAN-service-profile-level VMAC	disable
VMAC aging mode	mac-learning (the common mode)
VMAC control-protocol IPv6oE	disable
VMAC address count per xDSL/P2P port	32
VMAC address count per ONT	8
PPPoE/PPPoA MAC mode	**multi-mac**

Procedure

Configure the system ID. The system ID identifies an MA5600T/MA5603T/MA5608T in a network.

NOTICE

Make sure that you configure the system ID according to a data plan and that the system ID is unique in a network.

- In xPON access mode, run the **vmac olt-id** command to configure an OLT ID as the system ID.
- In access modes other than xPON, run the **vmac dslam-id** command to configure a DSLAM ID as the system ID.

NOTE

If both DSLAM ID and OLT ID exist in the system:

- Configure OLT ID as system ID for xPON access users.
- Configure DSLAM ID as system ID for non-xPON access users.

Step 1 Configure the maximum VMAC address count on each OLT port.

You can configure the maximum VMAC address count on each OLT port to limit the user count of the port. This operation prevents an excessive user count that will burden the system.

Run the **vmac port-vmac-count** command to configure the maximum VMAC address count on each OLT port.

Step 2 (Optional) In xPON access mode, configure the maximum VMAC address count on each ONT connected to the OLT.

You can configure the maximum VMAC address count on each ONT to limit the user count of the ONT. This operation prevents an excessive user count on a single ONT. An ONT with too many users will cause VMAC address insufficiency to the other ONTs connected to the same OLT port.

Run the **vmac ont-vmac-count** command to configure the maximum VMAC address count on each ONT.

Step 3 (Optional) Configure the PPPoE/PPPoA MAC address allocation mode to multi-MAC mode (1:1 VMAC mode).

NOTE

By default, the MAC address allocation mode is **multi-mac**. Perform this step only when the MAC address allocation mode has been modified to other values.

- You can run the **display pppoe mac-mode** command to query the current PPPoE MAC address allocation mode.
- You can run the **display pppoa mac-mode** command to query the current PPPoA MAC address allocation mode.

The MAC address allocation mode can be configured at system level (effective on users of all VLANs) or at VLAN service profile level (effective on users of the specified VLANs), depending on service deployment.

- For PPPoE users, run the **pppoe mac-mode** command to set the MAC address allocation mode to **multi-mac**.
- For PPPoA users, run the **pppoa mac-mode** command to set the MAC address allocation mode to **multi-mac**.

Step 4 Configure the VMAC aging mode.

The system ages unused VMAC addresses according to the aging mode in order to release VMAC address space. You can change the VMAC aging mode when the system default does not meet requirements.

In VLAN service profile mode, run the **vmac aging mode** command to configure the VMAC aging mode.

- **mac-learning**:

  The common aging mode. You can run the **mac-address timer** command to set the aging time. The system periodically checks for packets. If the system does not detect any packet, whether sent or received, carrying the VMAC address within twice the configured aging time, the system automatically releases the VMAC address, and this address can be allocated to another user.

- **dhcp**:

  In this mode, a user's VMAC address will not age if the user's IP address is allocated using DHCP. The VMAC ages only when the user's IP address is released or is not renewed after the lease expires.

  The DHCP aging mode applies only to the DHCP dialup service with the **multi-mac** MAC address allocation mode. This aging mode avoids frequent changes of system

entries because it maintains the mapping between the DHCP user's IP address and VMAC address before the IP address is released. For example, after a computer wakes up from hibernation, the computer will not perform DHCP dialup again. At this moment, the IP address allocated to the computer through DHCP is not released, so the computer can still use the mapping VMAC address.

Step 5 Enable the VMAC function.

NOTICE

The VMAC function can be configured at two levels: system level and VLAN service profile level. The VMAC function takes effect only when it is enabled at both levels.

- **System level**: Run the **vmac { enable | disable }** command to configure VMAC.
- **VLAN service profile level**: In VLAN service profile mode, run the **vmac { enable | disable } [ipoe | pppoe | pppoa] *** command to configure VMAC. You can enable or disable VMAC for all types of packets at VLAN level by running the **vmac { enable | disable }** command, or enable or disable VMAC for a specific packet type (PPPoE, PPPoA, IPoE) at VLAN level. Then bind this VLAN service profile to the VLAN.

Step 6 (Optional) Configure VMAC for IPv6oE protocol packets.

Perform this step only when using the VMAC feature in an IPv6 network.

Run the **vmac control-protocol ipv6oe** command to configure VMAC for IPv6oE protocol packets.

----End

Example

In an IPv4 network, all users use DHCP dialup for VDSL access. VLAN 10 requires the VMAC function. In such a service scenario, to set DSLAM ID to 0x0e02, maximum VMAC address count on each OLT port to 8, and VMAC aging mode to DHCP mode, do as follows:

```
huawei(config)#vmac dslam-id 0x0e02
huawei(config)#vmac port-vmac-count 8
huawei(config)#vlan service-profile profile-id 10
huawei(config-vlan-srvprof-10)#vmac aging-mode dhcp
huawei(config-vlan-srvprof-10)#vmac enable
huawei(config-vlan-srvprof-10)#commit
huawei(config-vlan-srvprof-10)#quit
huawei(config)#vlan bind service-profile 10 profile-id 10
huawei(config)#vmac enable
```

28.7.6 Configuring N:1 VMAC

In N:1 VMAC, the system generates one trusted virtual MAC addresses (VMAC) to replace source MAC addresses of end-users. N:1 VMAC prevents users that have untrustworthy MAC addresses from accessing carriers' networks and is an effective countermeasure to MAC spoofing. N:1 VMAC allows multiple MAC addresses to be translated into one VMAC address, so the access node only needs to record one VMAC entry, instead of multiple MAC entries, for a specified number of users. Using N:1 VMAC, an access node can accommodate more users by using the same MAC address space.

Prerequisites

The VMAC function of all packets, except PPPoA packets, conflicts with anti-MAC spoofing. When VMAC is enabled for all packets (except PPPoA), make sure that anti-MAC spoofing is not enabled at the same time. You can run the **display security config** command to query the status of anti-MAC spoofing.

Context

Table 28-14 System defaults of N:1 VMAC parameters

Parameter	Default Value
VMAC state	disable
VLAN-service-profile-level VMAC	disable
Maximum PPPoE session count of a port	8
Maximum PPPoE session count of each service port	This parameter does not have a default value.
Number of MAC addresses in each MAC address pool	256
MAC address allocation mode of PPPoE/PPPoA users	**multi-mac**

Procedure

Configure the maximum PPPoE session count.

You can configure the maximum PPPoE session count to limit the access user count. This operation prevents an excessive user count that will burden the system. The system supports two levels for configuring the maximum PPPoE session count: physical port level and service port level.

- Run the **pppoe max-session-count** command to configure the maximum PPPoE session count of a physical port on the OLT.
- Run the **pppoe max-session-count service-port** command to configure the maximum PPPoE session count of a service port that is configured on a physical port.

Step 1 (Optional) Configure a VMAC address pool for an xPON protection group.

NOTE
Perform this step only when an xPON protection group has been configured.

The system configured with xPON protection groups requires a VMAC address pool. The system performs MAC address translation (MAT) on idle MAC addresses in the VMAC address pool after receiving user packets. For details on related principles, see 28.7.3 N:1 VMAC Principles.

1. Configure a MAC address pool.

   Run the **mac-pool** [*pool-index*] **single-mac***startmac* [*scope*] command to configure the MAC address pool that will be used for replacing user MAC addresses through N:1 VMAC.

NOTICE

- For network security considerations, ensure that you have planned the MAC address pool during data planning, and that the MAC address pool does not conflict with the MAC addresses of other devices in the network.
- You can add or delete a MAC address pool, but cannot modify it.

2. Bind the MAC address pool to the xPON protection group.

   In protect-group mode, run the **bind mac-pool single-mac** command to bind the MAC address pool to the xPON protection group.

Step 2 Configure the PPPoE/PPPoA MAC address allocation mode to single-MAC mode (N:1 VMAC mode).

The MAC address allocation mode can be configured at system level (effective on users of all VLANs) or at VLAN level (effective on users of the specified VLAN), depending on service deployment.

- For PPPoE users, run the **pppoe mac-mode** command to set the MAC address allocation mode to **single-mac** for the system or for a VLAN service profile, or run the **pppoe vlan** command to set the MAC address allocation mode to **single-mac** for a single VLAN.
- For PPPoA users, run the **pppoa mac-mode** command to set the MAC address allocation mode to **single-mac**.

----End

Example

A PPPoE user is configured on service port 10 and locates in VLAN 10. To set the maximum PPPoE session count to 5 and set the MAC address allocation mode to **single-mac** for this user, do as follows:

```
huawei(config)#pppoe max-session-count service-port 10 5
huawei(config)#pppoe vlan 10 mac-mode single-mac
```

Assume that the system is configured with xPON protection group 0. To set the following parameters:

- Maximum PPPoE session count of service port 10: 5
- VMAC address pool, containing 10 MAC addresses, with start MAC address 0011-2222-3333
- MAC address allocation mode of VLAN 10: **single-mac**

do as follows:

```
huawei(config)#pppoe max-session-count service-port 10 5
huawei(config)#mac-pool single-mac 0011-2222-3333 10
huawei(config)#protect-group 0
huawei(protect-group-0)#bind mac-pool single-mac
huawei(protect-group-0)#quit
huawei(config)#vlan service-profile profile-id 10
huawei(config-vlan-srvprof-10)#pppoe mac-mode single-mac
huawei(config-vlan-srvprof-10)#commit
huawei(config-vlan-srvprof-10)#quit
huawei(config)#vlan bind service-profile 10 profile-id 10
```

29 Line Test

About This Chapter

Line tests are very helpful in diagnosing xDSL service faults.

The following lists the elements that cause xDSL service faults.

- Line-related elements. For example, the line is of poor contact and the line is not configured with a splitter.
- External noise. For example, radio frequency interference (RFI) or impulsive noise is detected.
- User-side faults. For example, the modem is not powered on and network port is not connected to a cable.

According to the statistics, only 10% of xDSL service faults are caused by device-related failures. A large percent of xDSL service faults are caused by line-related failures. Therefore, line tests are very helpful in diagnosing xDSL service faults.

29.1 SELT Test

Single ended loop testing (SELT) is a DSL line test initiated by a CO device and requiring no coordination of a CPE. It is used to test the physical and system configuration parameters of each service port.

29.1.1 Introduction

Single ended loop testing (SELT) can be used to collect information, such as the line length, background noise, line attenuation, and maximum rate. With the collected information, users evaluate lines and locate line faults.

- Before service provisioning, a SELT test is used to evaluate the quality and the service bearer capability of the subscriber line for a DSL port. In this case, pay attention to the information, such as the line length, line status, maximum upstream/downstream rate, and inband noise.
- After a fault is reported, a SELT test provides the information for diagnosing the line fault. In this case, pay attention to the information, such as the line status and line length.

In a SELT test, the xDSL chip set transmits test signals. When the test signals pass through the place where the impedance is discontinuous, the test signals are reflected. The xDSL chip set receives and analyzes the reflected signals, and then obtains the line condition.

NOTE

A SELT test does not require the coordination of a CPE. The test distance is affected by line signal attenuation, and the SELT test signals pass through the double line length.

29.1.2 Configuration

Procedure

(Optional) In ADSL or VDSL mode, run the **set selt level** command to set the level of single ended loop testing (SELT).

Time required for a SELT test depends on the SELT level. A higher-level SELT test requires more time and provides more accurate test data. The default SELT level is LEVEL0 (indicates the highest speed).

Step 1 Start a SELT test on an xDSL port.

- In VDSL mode, run the **vdsl selt** *portid* command to start a SELT test on a VDSL2 port.
- In ADSL mode, run the **adsl selt** *portid* command to start a SELT test on an ADSL2 port.

Step 2 Wait until the SELT test is completed and the test result is reported.

You can query the test result:

- Run the **display vdsl selt** *frameid/slotid/portid* command to query the test result.
- Run the **display adsl selt** *frameid/slotid/portid* command to query the test result.

Step 3 Learn about the line capability by the test result.

----End

Result

The following is an example test result of a SELT test.

```
huawei(config-if-vdsl-0/1)#
  Port 0/1/0 SELT finish successfully
  Frame/Slot/Port  : vdsl 0/1/0
  Updated time    : 2011-01-03 07:47:34+08:00
  Line length     : 2.74 m (9 feet)   //Whether the line length exceeds the maximum
limit.
  Line termination : open             //Whether a terminal is connected or the terminal
is faulty.
  Line gauge      : 0.4 mm (26 awg)   //Whether the line gauge meets site requirements.
  ------------------------------------------------------------
  SNR Margin(dB)   Upstream Rate(kbps)   Downstream Rate(kbps)
  ------------------------------------------------------------
            0              68760                92460
            3              68760                92460
            6              68760                92460
            9              64176                86296
           12              59592                80132
```

```
------------------------------------------------------------
```

System Response

Parameter	Description
Frame/Slot/Port	Indicates the subrack ID, slot ID, and port ID.
Updated time	Indicates the updated time of the SELT data.
Line length	Indicates the length of the tested line.
Line termination	Indicates the line termination status. The options are open, short, and unknown. • open: No modem is connected on the peer side. • short: A modem is connected on the peer side. • unknown: The test is abnormal and therefore no data is generated.
Line gauge	Indicates the gauge of the tested line.
SNR Margin	Indicates the gauge of the tested line.
Upstream Rate	Indicates the signal to noise (SNR) margin. Check whether the line quality meets the requirements by comparing the actual SNR margin with that specified in the network plan.
Downstream Rate	Indicates the upstream rate of the tested line. Check whether the line rate meets the requirements by comparing the actual line rate with the that specified in the network plan.

29.1.3 Reference Standards and Protocols

- RFC1155, RFC1157, and RFC1213 SNMP V1 series standards
- RFC1905, RFC1906, RFC1907, and RFC1908 SNMP V2 series standards
- RFC2571, RFC2572, RFC2573, RFC2574, RFC2576, RFC2578, and RFC2579 SNMP V3 series standards
- RFC959 FTP standard

29.2 MELT Test

Metallic line testing (MELT) is an integrated digital multimeter (DMM) test solution provided by a DSLAM device. The MELT function is integrated in a board. It can test the physical characteristics of the copper line for each service port.

29.2.1 Introduction

Metallic line testing (MELT) detects faults on the copper twisted pair from a service board to an xDSL user terminal. These faults include grounded subscriber lines, shorted wires within the same twisted pair, wires touching a high-voltage power line. MELT tests facilitate fast fault locating and minimize customers' operating expense (OPEX).

A MELT test can:

- Test electrical parameters, such as the compound capacitance, pure capacitance, voltage, resistance, and current of xDSL lines.
- Detect the potential physical faults on xDSL lines.
- Identify the fault range (such as intra-office faults or inter-office faults) and maintenance owner,

The N2510 can be used for analyzing the test results and locating the fault.

Two applications are included.

- Testing the electrical parameters of the line (that is, electrical parameter test)
- Testing the voice signal receiving capability of the line (that is, search tone test)

Electrical Parameter Test

In a MELT test, the MELT test chip sends test signals to the target port to test related electrical parameters, and then the chip calculates the major physical parameters of the line. (The MELT test chip can be regarded as a multimeter that can test the voltage, resistance, and capacitance.)

NOTE
During the MELT test, service running is not affected.

Search Tone Test

A search tone test helps users pinpoint the subscriber line connected to a specific port. The device plays the search tone on a specific port. Connect a phone (or a headset) to a line. Pick up the phone (or headset). If you can hear the search tone on the line through the phone (or headset), the line is connected to the specific port.

When the search tone test is started at the 600-ohm load, the search tone shall be applied symmetrically between wires A and B. The frequency of the search tone shall be 800 Hz. The voltage of the search tone at xDSL ports is within 120 mV and 330 mV.

NOTE
During the search tone test, SHDSL services are affected.

29.2.2 Electrical Parameter Test

Procedure

(Optional) In test mode, run the **xdsl melt polarity** command to set the mode of a twisted pair.

During the metallic line testing (MELT), the device applies a voltage between wires A and B to perform the test. If there is a passive test termination (PPA) model integrated into the phone socket, the line is not reachable. After the mode of the twisted pair is set, the device can apply a reverse voltage between wires A and B. By default, the polarity of a twisted pair is set to normal, which means the polarity of wire A is negative and the polarity of wire B is positive.

Step 1 Run the **xdsl melt** *frameid/slotid/portid* command in test mode to start the electrical parameter test.

Step 2 Wait until the electrical parameter test is completed and the test result is reported. Then, run the **display xdsl melt data** *frameid/slotid/portid* command to query the test result.

Step 3 Determine the line condition based on the test result.

----End

Result

The following is an example test result of an electrical parameter test.

```
huawei(config-test)#display xdsl melt data
frameid/slotid/portid<s><Length 1-15> }:0/2/0
 Command:
        display xdsl melt data 0/2/0

 Tested port: 0/2/0
 Start time of the test: 2010-08-19 15:54:20+08:00
 End time of the test:  2010-08-19 15:54:40+08:00
 I--Invalid   V--Valid
 ------------------------------------------------------------------------------
 Test Item                                          Flag  Result
 ------------------------------------------------------------------------------
 Conclusion:
   Does PPA exist?                                   I    no
   Does signature exist?                             V    no
   Does hazardous voltage (AC) exist?                   V    no
   Does hazardous voltage (DC) exist?                   V    no
   Line open                                        V    yes
   Line short                                       V    no
   Is the phone in the off-hook state?                  V    no
   Significant terminal device capacitance detected?       V    no

 Main measurement results:
   a->ground AC voltage(V)                             V    0.010
   b->ground AC voltage(V)                             V    0.013
   a->b AC voltage(V)                                 V    0.012
   Frequency of a->ground foreign AC voltage(Hz)           I    0
   Frequency of b->ground foreign AC voltage(Hz)           I    0
   Frequency of a->b foreign AC voltage(Hz)              I    0
   a->ground DC voltage(V)                             V    -0.191
   b->ground DC voltage(V)                             V    -0.141
   a->b DC voltage(V)                                 V    -0.051
   a->ground resistance(Ohm)                           V    10000000
   b->ground resistance(Ohm)                           V    10000000
   a->b resistance(Ohm)                               V    10000000
   b->a resistance(Ohm)                               V    10000000
   a->ground capacitance(nF)                           V    0
   b->ground capacitance(nF)                           V    0
   a->b capacitance(nF)                               V    1
   Capacitance of signature or ringer(nF)                 I    0
   Resistance of signature or ringer(Ohm)                 I    0
   Vzener(V)                                         I    -31.144
   Rzener(Ohm)                                       I    10000000
   Rbat_a(Ohm)                                       I    10000000
   Rbat_b(Ohm)                                       I    10000000

 Measurement results for reference:
```

```
  Aggregate a->ground conductance(uS)                    V   0.0
  Aggregate a->ground susceptance(uS)                    V   0.0
  Aggregate b->ground conductance(uS)                    V   0.0
  Aggregate b->ground susceptance(uS)                    V   0.0
  Aggregate a->b conductance(uS)                        V   0.0
  Aggregate a->b susceptance(uS)                        V   0.0
  AC foreign current on a wire(uA)                      V   0
  AC foreign current on b wire(uA)                      V   1
  DC foreign current on a wire(uA)                      V   -8
  DC foreign current on b wire(uA)                      V   -6
  Aggregate a->b capacitance under low voltage ramps(nF)    V  28
  Aggregate a->b capacitance under high voltage ramps(nF)   V  31
  Aggregate a->b capacitance under low AC signal(nF)       V  27
  Resistance in series with a->ground capacitance(Ohm)     V  5010911
  Resistance in series with b->ground capacitance(Ohm)     V  0
  Resistance in series with a->b capacitance(Ohm)


Applied measurement parameters:
  AC measurement frequency(Hz)                            275
  a->ground highest measurement DC voltage(V)            V   -3.622
  a->ground lowest measurement DC voltage(V)             V   -53.032
  b->ground highest measurement DC voltage(V)            V   -3.680
  b->ground lowest measurement DC voltage(V)             V   -53.718
  a->b measurement DC voltage(V)                        V   46.790
  b->a measurement DC voltage(V)                        V   -46.626
  Current mapping the a-wire highest voltage for test(uA)   V   -49
  Current mapping the a-wire lowest voltage for test(uA)    V   -83
  Current mapping the b-wire highest voltage for test(uA)   V   80
  Current mapping the b-wire lowest voltage for test(uA)    V   116
  Current mapping the a->b voltage for test(uA)           I   32
  Current mapping the b->a voltage for test(uA)           I   0
  Amplitude of AC voltage for a->ground test(V)           V  7
  Amplitude of AC voltage for b->ground test(V)           V  7
  Amplitude of AC voltage for a->b test(V)
---------------------------------------------------------------------------
```

System Response

Parameter	Description
Tested port	Indicates the ID of the test port.
Test Item	Indicate the test item. Currently, the test items include the following: • Does PPA exist?: indicates whether the PPA circuit exists. • Does signature exist?: indicates whether the signature circuit can be detected. • Does hazardous voltage (AC) exist?: indicates whether the AC voltage exceeds the threshold . • Does hazardous voltage (DC) exist?: indicates whether the DC voltage exceeds the threshold.

Parameter	Description
	• Line open: indicates that the circuit line is in the open circuit state. • Line short: indicates that the line is short-circuited. • Is the phone in the off-hook state?: indicates whether the phone is in the off-hook state. • Significant terminal device capacitance detected?: indicates whether the significant terminal device capacitance is detected. • a->ground AC voltage(V): indicates the wire A-to-ground (A-G) AC voltage (V). • b->ground AC voltage(V): indicates the wire B-to-ground (B-G) AC voltage (V). • a->b AC voltage(V): indicates the AC voltage (V) between wires A and B. • Frequency of a->ground foreign AC voltage(Hz): indicates the frequency (Hz) of the external AC voltage coupled to wire A and ground. • Frequency of b->ground foreign AC voltage(Hz): indicates the frequency (Hz) of the external AC voltage coupled to wire B and ground. • Frequency of a->b foreign AC voltage(Hz): indicates the frequency (Hz) of the external AC voltage coupled to wires A and B. • a->ground DC voltage(V): indicates the A-G DC voltage (V). • b->ground DC voltage(V): indicates the B-G DC voltage (V). • a->b DC voltage(V): indicates the DC voltage (V) between wires A and B. • a->ground resistance(Ohm): indicates the A-G resistance (Ohm). • b->ground resistance(Ohm): indicates the B-G resistance (Ohm). • a->b resistance(Ohm): indicates the wire A-to-wire B (A-B) resistance (Ohm) • b->a resistance(Ohm): indicates the wire B-to-wire A (B-A) resistance (Ohm) • a->ground capacitance(nF): indicates the A-G capacitance (nF). • b->ground capacitance(nF): indicates the B-G capacitance (nF). • a->b capacitance(nF): indicates the capacitance (nF) between wires A and B. • Capacitance of signature or ringer(nF): indicates the capacitance (nF) in a signature or ringer

Parameter	Description
	circuit. • Resistance of signature or ringer(Ohm): indicates the resistance (ohm) in a signature or ringer circuit. • Vzener(V): indicates the Zener diode breakdown threshold (V) in the case of an off-hook phone. • Rzener(Ohm): indicates the DC resistance (ohm) between wires A and B in the case of an off-hook phone. • Rbat_a(Ohm): indicates the equivalent coupling resistance (ohm) of wire A in the case of external voltage. • Rbat_b(Ohm): indicates the equivalent coupling resistance (ohm) of wire B in the case of external voltage. • Aggregate a->ground conductance(uS): indicates the A-G conductance (μS). • Aggregate a->ground susceptance(uS): indicates the A-G susceptance (μS). • Aggregate b->ground conductance(uS): indicates the B-G conductance (μS). • Aggregate b->ground susceptance(uS): indicates the B-G susceptance (μS). • Aggregate a->b conductance(uS): indicates the conductance (μS) between wires A and B. • Aggregate a->b susceptance(uS): indicates the susceptance (μS) between wires A and B. • AC foreign current on a wire(uA): indicates the A-G AC current (μA). • AC foreign current on b wire(uA): indicates the B-G AC current (μA). • DC foreign current on a wire(uA): indicates the A-G DC current (μA). • DC foreign current on b wire(uA): indicates the B-G DC current (μA). • Aggregate a->b capacitance under low voltage ramps(nF): indicates the aggregate capacitance between wire A and wire B under the low voltage test (nF). • Aggregate a->b capacitance under high voltage ramps(nF): indicates the aggregate capacitance between wire A and wire B under the high voltage test (nF). • Aggregate a->b capacitance under low AC signal(nF): indicates the capacitance between wire A and wire B under the low AC voltage signal test (nF).

Parameter	Description
	• Resistance in series with a->ground capacitance (Ohm): indicates the resistance in series with the A-G capacitance (ohm). • Resistance in series with b->ground capacitance(Ohm): indicates the resistance in series with the A-B capacitance (ohm). • Resistance in series with a->b capacitance(Ohm): indicates the resistance in series with the A-B capacitance (ohm) • AC measurement frequency(Hz): indicates the signal frequency (Hz) used in the complex impedance test. • a->ground highest measurement DC voltage(V): indicates the A-G upper measurement DC voltage (V). • a->ground lowest measurement DC voltage(V): indicates the A-G lower measurement DC voltage (V). • b->ground highest measurement DC voltage(V): indicates the B-G upper measurement DC voltage (V). • b->ground lowest measurement DC voltage(V): indicates the B-G lower measurement DC voltage (V). • a->b measurement DC voltage(V): indicates the (A-B) measurement DC voltage (V). • b->a measurement DC voltage(V): indicates the (B-A) measurement DC voltage (V). • Current mapping the a-wire highest voltage for test(uA): indicates the current (μA) generated to the A-G upper measurement DC voltage. • Current mapping the a-wire lowest voltage for test(uA): indicates the current (μA) generated to the A-G lower measurement DC voltage. • Current mapping the b-wire highest voltage for test(uA): indicates the current (μA) generated to the B-G upper measurement DC voltage. • Current mapping the b-wire lowest voltage for test(uA): indicates the current (μA) generated to the B-G lower measurement DC voltage. • Current mapping the a->b voltage for test(uA): indicates the current (μA) generated to the A-B measurement DC voltage. • Current mapping the b->a voltage for test(uA): indicates the current (μA) generated to the B-A measurement DC voltage. • Amplitude of AC voltage for a->ground test(V): indicates the amplitude of AC voltage (V)

Parameter	Description
	generated for the A-G measurement. • Amplitude of AC voltage for b->ground test(V): indicates the amplitude of AC voltage (V) generated for the B-G measurement. • Amplitude of AC voltage for a->b test(V): indicates the amplitude of AC voltage (V) generated for the A-B measurement.
Flag	Indicates the valid flag. It can be set to **I** (invalid) or **V** (valid).
Result	Indicates the test result of the corresponding test item.

29.2.3 Search Tone Test

Procedure

Run the **xdsl melt searching-tone** command in test mode to start the search tone test.

Step 1 Check whether the search tone is normal at the other end of the line using a headset or a receiver.

----End

Result

The system does not display any message after the command is executed successfully. You can use a headset or receiver to check whether the search tone is normal based on the configured search tone parameters.

29.2.4 Reference Standards and Protocols

- RFC1155, RFC1157, and RFC1213 SNMP V1 series standards
- RFC1905, RFC1906, RFC1907, and RFC1908 SNMP V2 series standards
- RFC2571, RFC2572, RFC2573, RFC2574, RFC2576, RFC2578, and RFC2579 SNMP V3 series standards
- RFC959 FTP standard
- ITU-T Recommendation G.996.2

29.3 DSL F5 OAM Loopback

DSL F5 OAM loopback, carrier over xDSL lines, automatically checks the continuity of ATM links between user terminals and access devices.

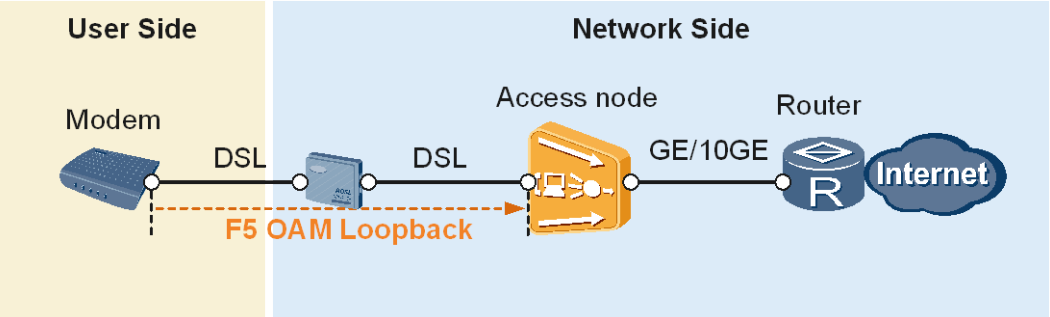
29.3.1 Introduction

An ATM network has five levels of OAM, F1 through F5. Among the five levels, F1 through F3 serve the physical layer, and F4 and F5 serve the ATM layer. At the ATM layer, F4 OAM operates at virtual path (VP) level and F5 OAM operates at virtual channel (VC) level. F5 OAM checks the continuity of an ATM link with a specified VPI/VCI and monitors the link performance. VPI is the abbreviated form of virtual path identifier and VCI is that of virtual channel identifier.

F5 OAM supports fault management, performance management, activation or deactivation, and system management. F5 OAM loopback is a means of fault management.

DSL F5 OAM loopback complies with ITU-T I.610. Figure 29-1 shows the typical application of DSL F5 OAM loopback on an access network.

Figure 29-1 Typical application of DSL F5 OAM loopback on an access network

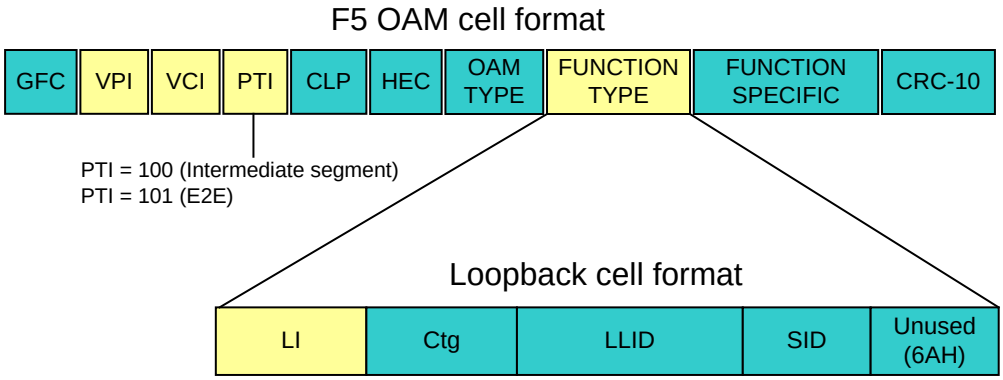


29.3.2 Principles

Cell Format

Figure 29-2 shows the format of a DSL F5 OAM loopback cell.

Figure 29-2 Format of a DSL F5 OAM loopback cell



- VPI/VCI field

According to ITU-T I.610, the continuity of a link with a specified VPI/VCI can be checked only if the following requirements are met:

- The VPI/VCI of F5 OAM cells on the access device is the same as that of common user cells on the terminal.
- Both the access device and the terminal support DSL F5 OAM loopback.

- PTI field
  - A payload type identifier (PTI) differentiates F5 OAM cells from common cells. The PTI field of an F5 OAM cell may take value 100 or 101. Value 100 indicates cells in intermediate segments and value 101 indicates end-to-end (E2E) cells.
  - The device where F5 OAM cells pass through on an ATM network automatically fills a value in the PTI field.
- FUNCTION TYPE field

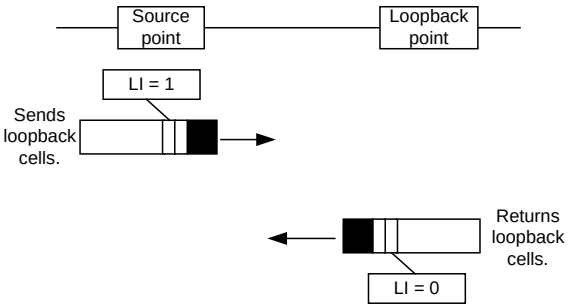
  DSL F5 OAM loopback cells are defined in the FUNCTION TYPE field. Figure 29-2 shows the cell format.

  In this field, loopback identifier (LI) indicates a request or response cell. The loopback source point fills the LI field with 00000001. That is, the LI field is set to 1. The loopback point then returns cells after changing the LI field to 00000000. That is, the LI field is set to 0.

Loopback Process

DSL F5 OAM loopback supports loopback detection at two ends. Both an access device and a terminal can initiate loopback detection. Figure 29-3 shows the DSL F5 OAM loopback process. The end that initiates loopback detection is the source point and the end from which the loopback cells are returned is the loopback point.

Figure 29-3 DSL F5 OAM loopback process



The loopback process is as follows:

1. The source point sets the LI field to 1 to initiate a link continuity check.
2. The loopback point checks whether loopback cells from the source point are received.
   - If the loopback cells are received, go to step 3.
   - If no loopback cells are received, go to step 5.
3. The loopback point checks whether continuity detection is enabled based on services.
   - If continuity detection is enabled, go to step 4.
   - If continuity detection is disabled, the loopback point sets the LI field to 0 and returns loopback cells. Then, go to step 5.

4. The loopback point checks whether the VPI/VCI of the cells sent from the source point is the same as that of existing cells configured on the loopback point.
   - If the VPIs/VCIs are the same, the loopback point sets the LI field to 0 and returns the loopback cells. Then, go to step 5.
   - If the VPIs/VCIs are different, the loopback point discards the cells. Then, go to step 5.
5. The source point determines the link continuity 5s after initiating the link continuity check.
   - If the source point receives response cells, the loopback detection is successful, and the link and service are functional.
   - If the source point does not receive any response cells, the loopback detection fails, and the link or service is malfunctioning.

NOTE

- Loopback detection checks the continuity of a link at a moment. The source point determines the number of consecutive detection attempts and the waiting duration for a response.
- If the loopback point supports service-based loopback detection, it can check service status.

Cell Processing Policy

Table 29-1 lists the settings of the functions supported by an access device working as a loopback point and the processing policies for loopback cells.

Table 29-1 Settings of the functions supported by an access device working as a loopback point and the processing policies for loopback cells

Function Setting	Command	Processing Policy for Loopback Cells
Disabling DSL F5 OAM loopback	**f5 oam loopback disable**	Discards loopback cells.
Enabling F5 OAM physical link detection and disabling service-based continuity detection	**f5 oam loopback enable uncheck**	Sets the LI field to 0 and returns loopback cells. In this case, the physical link between the terminal and the access device is functional.
Enabling service-based continuity check	**f5 oam loopback enable check**	Sets the LI field to 0 and returns loopback cells. In this case, the physical link and service flows between the terminal and the access device are functional.

29.3.3 Application

DSL F5 OAM loopback enables an access device to support smooth Internet access mode switching from PPP to Dynamic Host Configuration Protocol (DHCP) for users connected to a modem that supports special functions.

This section uses an example to describe how to implement DSL F5 OAM loopback on an access device.

Service Requirements

To reduce operating expense (OPEX), the carrier intends to smoothly switch the Internet access mode from PPP to DHCP for users connected to the modems of type 2 without affecting existing broadband services.

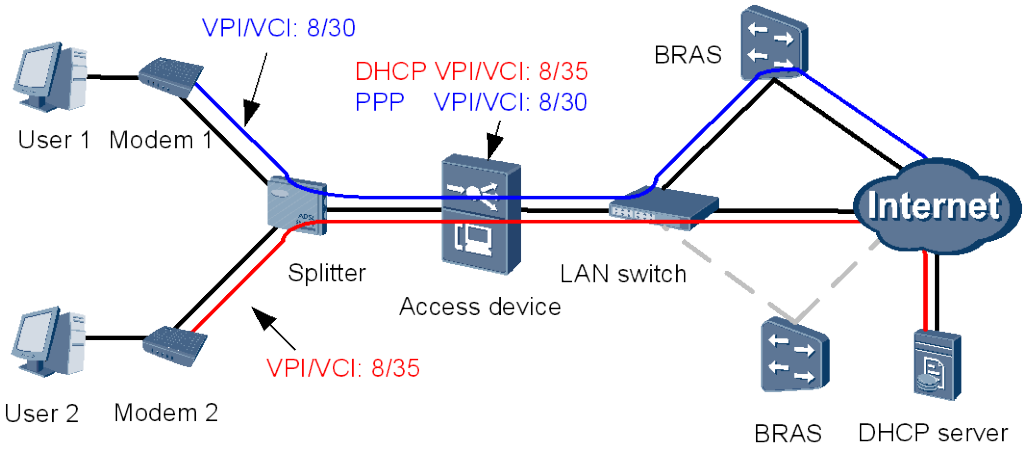
Two types of modems are available on a carrier's network:

- The modems of type 1 provide common functions and support the Internet access only in PPP mode.
- The modems of type 2 provide the following special functions:
  - Initiate loopback detection only in DHCP mode.
  - Support automatic switching from PPP to DHCP.

Networking Diagram

Figure 29-4 shows the networking diagram for Internet access in PPP and DHCP modes.

Figure 29-4 Networking diagram for Internet access in PPP and DHCP modes



Prerequisites

- The VPI/VCI for the Internet access in PPP mode is 8/30 and that for the Internet access in DHCP mode is 8/35.
- Users 1 and 2 can access the Internet.
- Modem 1 provides common functions and supports the Internet access only in PPP mode.
- Modem 2 provides the following special functions:
  - Initiates loopback detection only in DHCP mode.
  - Supports automatic switching from PPP to DHCP.

Device Configuration

Run the following commands to enable service-based DSL F5 OAM loopback:

```
huawei(config)#f5
{ oam<K> }:oam
```

```
{ loopback<K> }:loopback
{ disable<K>|enable<K> }:enable
{ check<K>|uncheck<K> }:check

 Command:
      f5 oam loopback enable check
```

Service Deployment

For user 1:

User 1 connects to modem 1. After the VPI/VCI 8/30 is configured on the access device, user 1 can access the Internet in PPP mode.

For user 2:

User 2 connects to modem 2. User 2 accesses the Internet in PPP mode. After DHCP services are deployed on the network:

1. Modem 2 automatically initiates link detection in DHCP mode.
2. After receiving loopback cells sent from modem 2, the access device checks whether the VPI/VCI of the cells sent from modem 2 is the same as that of existing service cells.
   - If the VPIs/VCIs are the same, the access device returns loopback cells. Modem 2 automatically switches the Internet access mode to DHCP upon receiving the response from the access device.
   - If the VPIs/VCIs are different, the access device does not respond to the request of modem 2. If modem 2 does not receive a response for three consecutive attempts, the PPP Internet access mode of user 2 remains unchanged.

30 Power Saving and Maintenance

About This Chapter

This topic describes the power saving feature of the system from two aspects: stepless speed adjustment of the fan and power cutoff of the board. It also describes the maintenance feature of the system from two aspects: power cutoff of the board and recording the model and running information for the fan and power module.

30.1 Overview of the Power Saving and Maintenance Feature

The power saving feature is related to the following two items:

- The fan rotating speed is precisely controlled based on the temperature on the parts (boards), and the fans do not rotate at a constant speed, thus reducing the power consumption.
- The device supports five power saving features to reduce power consumption. The power saving features are automatic unbinding for an extended profile, universal power saving, innovate power saving, automatic board power off, and vectoring power saving.

The maintenance feature is related to the following two items:

- The remote board can be manually powered off and then powered on for maintenance, which is similar to the hot plug operation.
- The model and running information of the fan and power module can be recorded, thus reducing the process cost for preparing the spare parts of multiple versions.

30.2 Power Saving

This topic describes the power saving feature of the system which is implemented by means of stepless speed adjustment of the fan and power cutoff of a service board.

30.2.1 Introduction

Definition

- The fans for the subrack of the MA5600T/MA5603T/MA5608T do not rotate at a constant speed. They automatically implement the stepless speed adjustment according to the temperature inside the subrack detected by the temperature sensor, thus reducing the power consumption.
- The MA5600T/MA5603T/MA5608T supports five power saving features to reduce power consumption. The power saving features are automatic unbinding for an extended profile, universal power saving, innovate power saving, automatic board power off, and vectoring power saving.

Purpose

The purpose of the power saving feature is to reduce the power consumption and heat consumption of the system.

Benefits

- The power-saving feature helps reduce system power consumption efficiently and therefore reduces carriers operating expense.
- The formula for calculating the power-saving amount is as follows:
  - Power-saving amount (broadband service board) = Total power-off time of the board (hours) x Static power consumption of the board (watt)
  - For example, the total power-off time of a broadband service board is 720 hours, and the static power consumption of the board is 55 W. The total power-saving amount of the board is 39.6 kWH (720 hours x 55 W).

30.2.2 Principle

This topic describes the working principle of the power saving feature.

Stepless Speed Adjustment of the fans on the Shelf

The working principle of the automatic stepless speed adjustment of the fans is as follows:

1. The system queries the temperature on all the boards in the subrack every 10 minutes, and directly omits the boards that do not support querying the temperature.
2. If the queried temperature meets the speed adjustment condition, the system issues the rotating speed adjustment command and the corresponding duty ratio to the fans on the subrack to adjust their rotating speed. Table 30-1 shows the mapping between the speed adjustment of the fans and the temperature control point of the boards.

Table 30-1 Mapping between the speed adjustment of the fans and the temperature control point of the boards

Control Point	Action
Tmin	If the temperature on all the boards is less than Tmin, the rotating speed of the fans on the subrack is decreased by N% (N = 10).
Tmax	If the temperature on a board is larger than Tmax, the rotating speed of the fans on the subrack is increased by N% (N = 10).
Tminor	If the temperature on a board is larger than Tminor, the green LED is on for 1s and off for 1s repeatedly, and the rotating speed of the fans on the subrack is increased to the full speed.
Tmajor	If the temperature on a board is larger than Tmajor, the orange LED is on for 0.25s and off for 0.25s repeatedly, the rotating speed of the fans on the subrack is increased to the full speed, and the system generates a high-temperature alarm. The user needs to take power saving measures, such as shutting down the port on the board or cutting off the power of the board.
Tcritical	If the temperature on a board is larger than Tcritical, the system cuts off the power of the board (excluding the control board) and generates a high-temperature recovery alarm. The cutoff time is 15 minutes. After that, the system forcibly powers on the board, and adjusts the rotating speed of the fans according to the temperature on other boards.

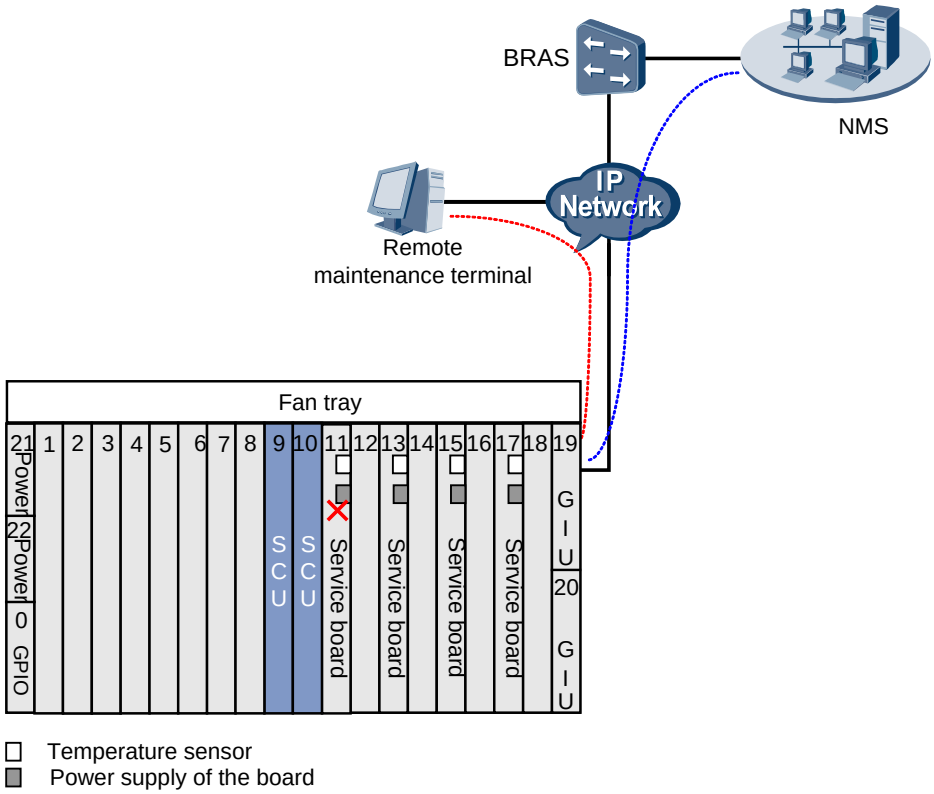
3. The rotating speed of the fans on the subrack is adjusted to the expected value.

Automatic/Manual Power Cutoff of the Board

When the temperature on a board reaches the threshold of danger or the board is not configured with any service, measures should be taken to reduce the power consumption. Currently, the power of the board is cut off to reduce the power consumption.

Figure 30-1 shows the power saving principle of the automatic/manual power cutoff of the board.

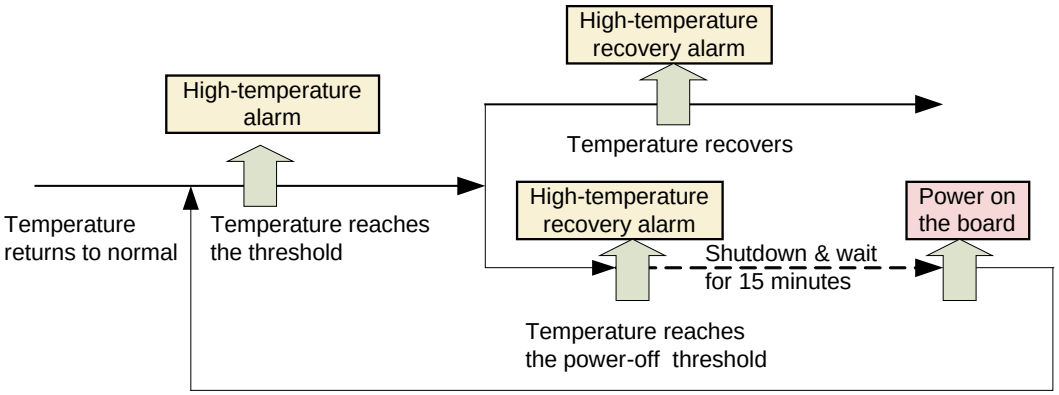
Figure 30-1 Power saving principle of thee automatic/manual power cutoff of the board



The power saving principle of the automatic/manual power cutoff of the board is as follows:

1. The sysetm polls the temperature every 20s. When detecting that the temperature on a board exceeds the temperature threshold, the system automatically cuts off the power supply of the board, and then powers it on in 15 minutes. The processing is shown as Figure 30-2.
2. When the system detects that the temperature on a board is too high or the port on the board is not configured with any service, you can manually cut off the power of the board through the CLI or NMS to reduce the power consumption. After powering off the board, you can power it on again through the CLI or NMS.
   - In case of high temperature, when you run a command to manually cut off the power of a board that is providing a service, the system prompts that a service is running on the board and the service will be interrupted if the power is cut off. You need to determine whether to cut off the power of the board.
   - In the case that the port on a board is not configured with any service, the system does not display any message when you run a command to manually cut off the power of the board.
   - When a command is issued from the NMS to cut off the power of a board, the system directly cuts off the power of the board in any case.

Figure 30-2 Processing of automatic power cutoff of the board



Refined Power Saving for Boards

Table 30-2 Refined power saving for boards

Power Saving Feature	Principle	Remarks
Automatic unbinding for an extended profile	When a line profile bound to asymmetric digital subscriber line (ADSL2+) service ports is changed, the system automatically unbinds the extended profile to prevent the power saving configuration in the extended profile from affecting services on the service ports.	These five features can be implemented only through a professional service.
Universal power saving	This feature involves metallic loop test (MELT) and control board power saving. • MELT power saving: When no MELT test is performed within 15 minutes, the MELT module goes to the sleep state. When a MELT test is performed, the MELT module immediately restores to the work state. • Control board power saving: The system monitors and manages various interface circuits on the control board. Specifically, the system automatically shuts down idle interface circuits. When the interface circuits are to be used, the system automatically starts them so that they restore to the normal working state.	
Innovative power saving	This feature is supported by line drivers (LDs). The system determines the power saving status for LDs according to the profile configuration for each service port on a very-high-speed digital subscriber line 2 (VDSL2) board. This ensures the lowest	

Power Saving Feature	Principle	Remarks
	power consumption for LDs.	
Automatic board shut down or sleep mode	The system automatically identifies boards and powers off them or make them to go to sleep mode. When the boards are to be used, the system automatically powers on or makes them to go to work state.	
Power saving of vectoring processing boards	The system automatically identifies idle vectoring processing boards and powers off them or make them to go to sleep mode. When the boards are to be used, the system automatically powers on or makes them to go to work state. In addition, the system monitors and manages various interface circuits on the vectoring processing boards. Specifically, the system automatically shuts down idle interface circuits. When the interface circuits are to be used, the system automatically starts them so that they restore to the normal working state.	

30.3 Maintenance

Describe the maintenance feature and the principle of this feature.

30.3.1 Introduction

Definition

- The MA5600T/MA5603T/MA5608T supports power cutoff of the board, and can be forcibly powered off/on even when it is down or faulty. This is like remote hot plug of the board, meeting the requirement of remote maintenance.
- The MA5600T/MA5603T/MA5608T supports recording the model and running information for the fans on the subrack and the power module. When a fan on the subrack or the power module is faulty, the model and running information of the faulty part can be queried. In this manner, the maintenance engineer can bring the correct spare parts to the field and analyze the cause of the fault according to the running information.

Purpose

The purpose of the maintenance feature is to reduce the human cost for multiple site visits of the maintenance engineer and the process cost of preparing spare parts of multiple versions.

- The maintenance engineer can cut off the power of the faulty board through the CLI or NMS in the CO instead of removing and inserting the board on site, thus recovering the service.

- When a fan on the subrack or the power module is faulty, the maintenance engineer can query the information about the faulty part and the running information in the last three times through the CLI in the CO, thus preparing the correct spare parts, reducing the cost of preparing spare parts of multiple versions, and analyzing the cause of the fault according to the running information.

Specifications

None

Limitations

- The control board, power board, and GIU upstream board do not support power cutoff of the board.

30.3.2 Principle

Manual Power Cutoff of the Board

The maintenance principle of the manual power cutoff of the board is as follows:

1. When a board is down or faulty:
   - Before: The maintenance engineer needs to go to the site and power on the board again on site by removing and inserting the board.
   - Now: The maintenance engineer needs not go to the site. Instead, the maintenance engineer can power on the board again by forcibly cutting off the power of the board through the CLI or NMS in the CO.
2. The manual power cutoff of the board reduces the human cost on multiple site visits of the maintenance engineer.

Recording the Model and Running Information for the Fans on the Shelf and the Power Module

The current status analysis for the maintenance of the fans on the subrack and the power module is as follows:

- When a fan on the subrack or the power module is faulty, the model and running information of the faulty module are not recorded on the host.
- The fans on the subrack and the power module have multiple models, and the maintenance engineer needs to prepare multiple types of modules for the site, which increases the preparation cost and process cost.
- The fans on the subrack and the power module in a site become faulty for multiple times, but there is no corresponding information for fault analysis.

The system records the model and running information for the fans on the subrack and the power module as follows:

1. When the fans on the subrack and the power module work in the normal state, the system records the current system time and the running information in the last three times. The following points are included:
   - System time (Systime)
   - EMU type (EMU type)
   - EMU name (EMU name)

- Fan type (FAN type)
- Software version (Soft ver)

2. When a fan on the subrack or the power module is faulty, you can run a command to query the detailed information about the fan or the power module.
   - The model of the fan or power module to be replaced can be precisely recognized, which effectively saves the preparation cost of the maintenance engineer.
   - The running information can be used to analyze the fault and find the root cause, thus reducing the fault possibility of the fan and the power module.

31 Ethernet OAM

About This Chapter

Ethernet Operation, Administration, and Maintenance (OAM) is used to operate, administrate, and maintain Ethernet networks. It covers three features: connectivity fault management (CFM), Ethernet in the first mile (EFM), and performance monitoring (PM).

31.1 Ethernet OAM Introduction

Feature Overview

Ethernet is a widely used local area network (LAN) technology featuring rich bandwidth and low cost. As the application of Ethernet is spanning from carrier networks to metropolitan area networks (MANs) and wide area networks (WANs), network administration and maintenance become increasingly important. However, traditional Ethernet is unable to provide end-to-end (E2E) service management, fault detection, or performance monitoring, which are instead implemented by IP-based mechanisms. These mechanisms (such as ping messages) at the IP layer are used to operate and manage the entire network, but they can not meet operation and maintenance requirements of the Ethernet bottom layer that has no IP addresses. Against this backdrop, the IEEE and ITU-T have done a lot of researches and worked out their own Ethernet OAM standards.

Ethernet Operation, Administration, and Maintenance (OAM) covers three features: connectivity fault management (CFM), Ethernet in the first mile (EFM), and performance monitoring (PM). Table 31-1 describes the main functions of these features.

Table 31-1 Ethernet OAM functions and application scenarios

Feature	Description	Application Scenario
CFM	CFM is defined in the IEEE 802.1ag (802.1ag for short). It is an OAM feature for detecting and locating Ethernet connectivity faults by including the functions of continuity check (CC), loopback (LB), and linktrace (LT). The ITU-T Y.1731 (Y.1731 for	CFM mainly applies to access and convergence nodes. It is used to monitor connectivity on the entire network and to locate connectivity faults for end-to-end (E2E) Ethernet links. For example, in IPTV services, CFM can be used to monitor network connectivity between a user terminal

Feature	Description	Application Scenario
	short) covers all the CFM functions defined in the 802.1ag and adds more information types to enhance the CFM capability.	and the service provider's IPTV server and to determine the location of a fault.
EFM	EFM is defined in the IEEE 802.3ah (802.3ah for short). It is an OAM feature for detecting quality and connectivity of last-mile Ethernet links.	EFM packets cannot be forwarded across multiple hops. Therefore, EFM mainly applies to the user access network (that is, the last mile) and also to the Ethernet physical link that directly connects two devices. For example, EFM can be used to monitor the link between a user access board and a user terminal and to notify the system of link exceptions using Ethernet OAM link events. Link exceptions can be link interruptions or that the number of received error packets exceeds the preset threshold due to deteriorating link quality.
PM	PM is defined in the Y.1731. It is an OAM feature for measuring and testing Ethernet performance, including loss measurement (LM), delay measurement (DM), and throughput measurement.	PM is mainly used to evaluate the network performance and quality and to discover latent network defects (such as an increasing packet loss rate due to line aging), reducing user complaints and improving quality of service (QoS).

31.2 Reference Standards and Protocols

Table 31-2 Reference standards and protocols of the Ethernet OAM feature

Ethernet OAM Type	Reference Standards and Protocols
CFM	• ITU-T Y.1731: OAM functions and mechanisms for Ethernet based networks • IEEE 802.1ag-2007 VLAN Amendment: 5 Connectivity Fault Management
EFM	IEEE 802.3ah: Operations, Administration, and Maintenance (OAM)
PM	ITU-T Y.1731: OAM functions and mechanisms for Ethernet based networks

NOTE

Connectivity fault management (CFM) is implemented according to the formal IEEE 802.1ag-2007 and is not compatible with Draft6.0.

ITU-T Y.1731 mentioned in this document is of version 07/2011.

31.3 Differences in Implementing Y.1731 and 802.1ag on Access Device

The ETH OAM of the access device is implemented based on concepts defined by 802.1ag. For concepts defined by both 802.1ag and Y.1731, there is no difference in implementation. For concepts defined by Y.1371 but not defined in 802.1ag, the similar concepts in 802.1ag are used. The following table describes the differences in implementing related concepts of Y.1731 and 802.1ag on the access device. The similar concepts of 802.1ag and Y.1731 are in the same row.

Concept of 802.1ag	Concept of Y.1731	Implementation of Differences Between Y.1731 and 802.1ag
Maintenance domain (MD)	No similar definition	For Y.1731, **MD name** is configured to **no name**.
Maintenance association (MA)	Maintenance entity group (MEG)	For Y.1731, ensure that the following requirements are met when configuring an MA. • **MA ID** is configured to **MEG ID**. • **MA name** must use the ICC-based format.
Maintenance point (MP)	MEG point (MP)	There is no difference in implementing MP in Y.1731 and 802.1ag.
Maintenance association end point (MEP)	MEG end point (MEP)	There is no difference in implementing MEP in Y.1731 and 802.1ag.
Maintenance association intermediate point (MIP)	MEG intermediate point (MIP)	There is no difference in implementing MIP in Y.1731 and 802.1ag.

31.4 CFM (802.1ag and Y.1731)

Connectivity fault management (CFM) is defined in the IEEE 802.1ag (802.1ag for short). It is an OAM feature for detecting and locating Ethernet connectivity faults. The ITU-T Y.1731 (Y.1731 for short) covers all the CFM functions defined in the 802.1ag and adds more information types to enhance the CFM capability.

31.4.1 CFM Introduction

The IEEE and ITU-T have worked out their own standards for monitoring Ethernet connectivity and identifying connectivity faults for E2E Ethernet links.

- The 802.1ag (also known as CFM) defines Ethernet Operations, Administration, and Maintenance (OAM) functions, which are used for detecting and identifying link connectivity faults on a network involving multiple independent maintenance organizations.
- The Y.1731 covers all functions defined in the 802.1ag and enhances some functions.

The access device supports CFM defined in the 802.1ag and Y.1731. Table 31-3 describes three functions of CFM.

Table 31-3 Three functions of CFM

Function	Description
CC Principle	Monitors connectivity of Ethernet links in real time and reports link exceptions by alarms.
LT Principle	• Obtains MAC addresses of intermediate devices along the Ethernet link between two devices. • Determines the location of a link fault.
LB Principle	Unicast LB • Checks connectivity of the Ethernet link between two devices. • Determines the location of a link fault. Multicast LB • Checks connectivity of Ethernet links between one device and multiple other devices. • Obtains the MAC address of a remote device that has the same maintenance level as the local device. Bidirectional diagnostics test: Checks the packet loss rate and bit error rate of the bandwidth-specific bidirectional Ethernet link between two devices to determine whether the bandwidth meets user requirements.
AIS Principles	• Suppress alarms. • Send link fault information to the peer with higher-level.
LLF Principles	Upon link fault detection, the device sends information to the peer end which is still communicated with the device. Then the peer end can react such as shut down the port automatically. LLF is often used in the FTTO scenarios of national broadband network. When the leased channel of the retail service provider (RSP) is faulty, LLF function makes it possible that the RSP and its users are informed by the operator immediately.

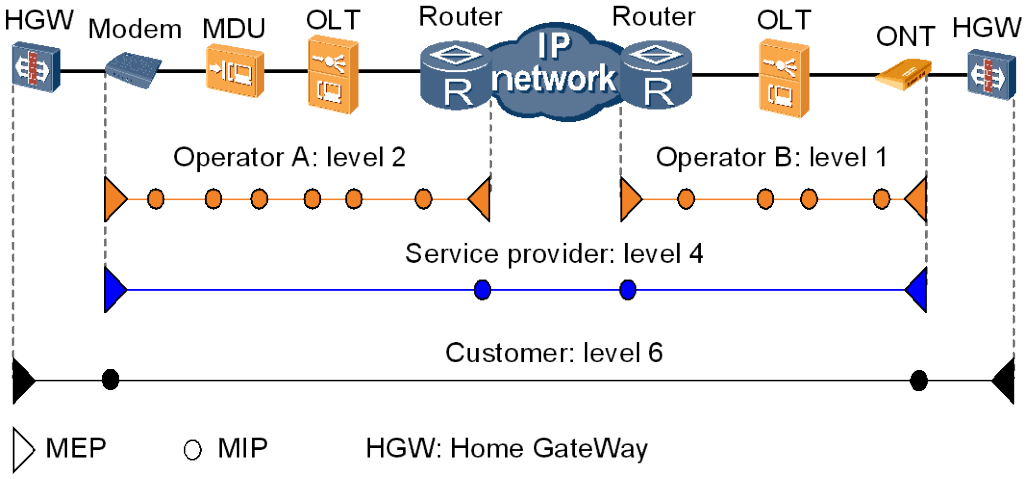
Y.1731 and 802.1ag have the following differences in implementing the CFM function.

- The continuity check (CC) function and principle defined by Y.1731 and 802.1ag are the same only with one difference in the format of the continuity check message (CCM). Because of the difference, there are some limitations in usage of the CFM function when the access device is interconnected with the peer end that only supports Y.1731.
  - When MD is being configured, configure **MD name** to no name.
  - When MA is being configured, MA ID is identical to MEG ID. **MA name** must use the ICC-based format.
- 802.1ag defines the unicast LB. Y.1731 defines the unicast LB, multicast LB, and bidirectional diagnostics test. In addition, Y.1731 provides the extended loopback message (LBM), and the size and sending interval of LBMs can be set and disorder in the loopback reply (LBR) message returned by the peer end can be detected.
  - No matter whether the peer end device supports Y.1731 or 802.1ag, when implementing an LB (including the unicast LB, multicast LB, and bidirectional diagnostics test), the access device can receive the LBR message returned by its peer end. In addition, the test result does not vary with the protocol used at the peer end.
  - No matter whether the peer end device supports Y.1731 or 802.1ag, the access device can correctly reply with the LBR messages.
- Y.1731 and 802.1ag have no difference in implementing the LT function. No matter whether the peer end device supports Y.1731 or 802.1ag, the access device can implement the link trace (LT) and correctly reply with the linktrace reply (LTR) messages.
- ETH-AIS is only defined in Y.1731.
- LLF is a function defined by Huawei. It is implemented based on the information format defined in 802.1ag.

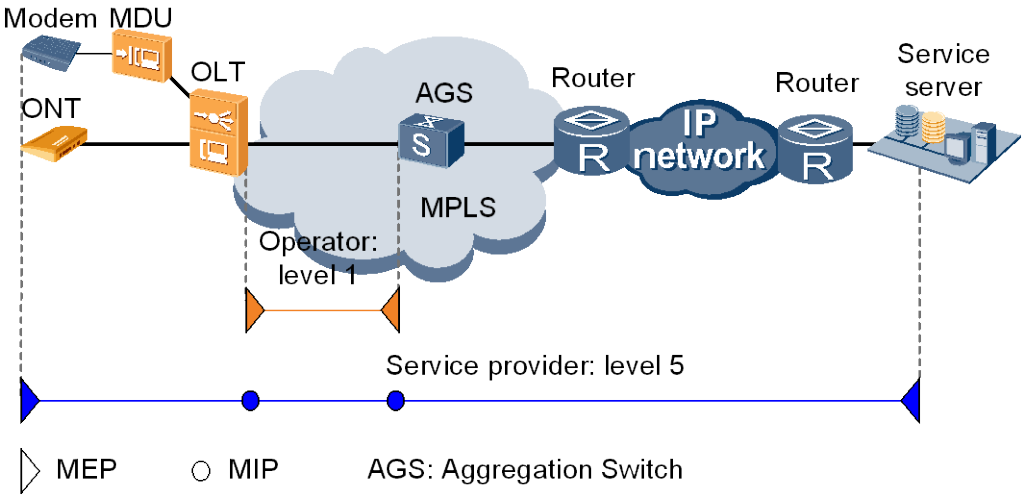
31.4.2 CFM Network Application

In carrier-class Ethernet networks, different organizations are required to provide different administration and maintenance scopes and contents. Therefore, carrier-class Ethernet needs to be maintained hierarchically. Three types of organizations are generally involved in carrier-class Ethernet services: customers (for example, private network users), service providers, and network operators. Customers purchase Ethernet services from service providers, service providers can use their own networks or leased networks to provide end-to-end (E2E) Ethernet services, while network operators provide networks for transmitting services.

The following picture shows a typical CFM network. In the figure, locations of maintenance association end points (MEPs) and maintenance domain intermediate points (MIPs) are marked, and different maintenance domains (MDs) are configured for three maintenance levels (customer, service provider, and network operator).



CFM also applies to MPLS networks for connectivity diagnosis, as shown in the picture below.



31.4.3 CFM Basic Concepts

This topic describes some basic concepts of the connectivity fault management (CFM), such as the maintenance domain (MD), maintenance association (MA), maintenance association end point (MEP) and maintenance association intermediate point (MIP) that we need to understand before getting to know the working principle of the CFM function.

MD

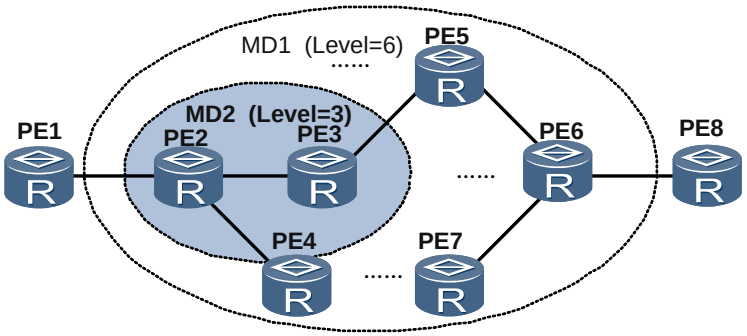
A maintenance domain (MD) is used to divide the network covered by the connectivity fault test.

In order to locate a fault precisely, the concept of **level** is introduced in MD. An MD provides eight levels, represented by the integer number 0-7. A larger number indicates a higher level and a larger-range MD. An MD can be embedded into or neighbor on another MD but an MD cannot cross another. A low-level MD is embedded into a high-level MD, but a high-level MD cannot be embedded into a low-level MD.

The MD level makes the fault locating precise and easy. As shown in Figure 31-1, MD2 is embedded into MD1. If MD1 detects a fault, it is suspected that the links or the devices from PE2 to PE6 are faulty. If MD2 does not detect any fault, PE2, PE3 and PE4 work properly. Therefore, the fault is narrowed down to the links between PE5, PE6, and PE7 or these devices.

In actual application, if an MD contains another small-range MD and the connectivity test is performed for the large-range MD, CFM packets need to traverse the small-range MD. To achieve that purpose, set the level of the large-range MD higher than that of the small-range MD. For example, in the network as shown in Figure 31-1, MD1 contains MD2 and CFM packets of MD1 need to traverse MD2. The level of MD1 is set to 6 and the level of MD2 is set to 3. In this way, CFM packets of MD1 can traverse MD2 and CFM for the entire MD1 can be implemented. Moreover, CFM packets of MD2 will not spread to MD1.

Figure 31-1 MD levels



PE: Provider Edge

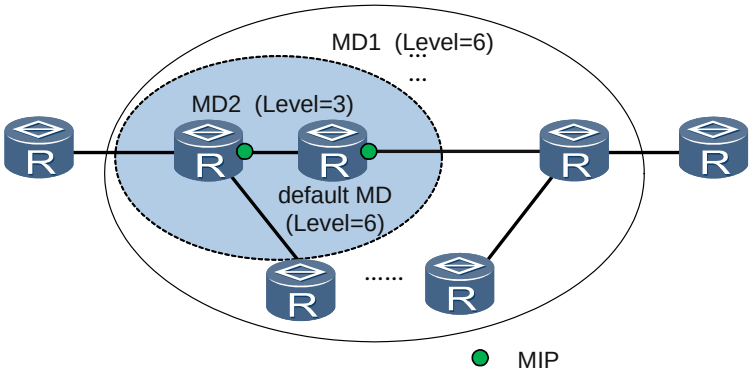
CFM packet interaction and basic CFM functions are implemented based on MDs. Proper MD planning can help network administrators locate faults quickly.

Default MD

The default MD is used for the high-level MD to sense the internal topology of the low-level MD.

As shown in Figure 31-2, a low-level MD is embedded into a high-level MD, and devices in the high-level MD may be the edge and intermediate devices in the low-level MD. When CFM protocol packets of the high-level MD traverse the low-level MD, the packets are transmitted transparently. If the default MD is not configured and the high-level MD needs to sense the internal topology of the low-level MD, a maintenance association intermediate point (MIP) with a specific level needs to be created on the specific port of the device in the low-level MD and the MIP is used to reply to devices in the high-level MD with loopback reply (LBR) or linktrace reply (LTR) messages.

Figure 31-2 Default MD



If the default MD with the same level as the high-level MD is configured on the devices in the low-level MD, MIPs of corresponding levels based on default MDs are generated to reply to devices in the high-level MD with LBR or LTR messages. In this way, the high-level MD can sense the topology changes in the low-level MD and CFM can be implemented in the entire MD1.

The level of the default MD is the same as that of the high-level MD but is higher than the levels of all MDs to which the MEPs configured on the device belong.

MA

An MA specifies the range a CCM can reach.

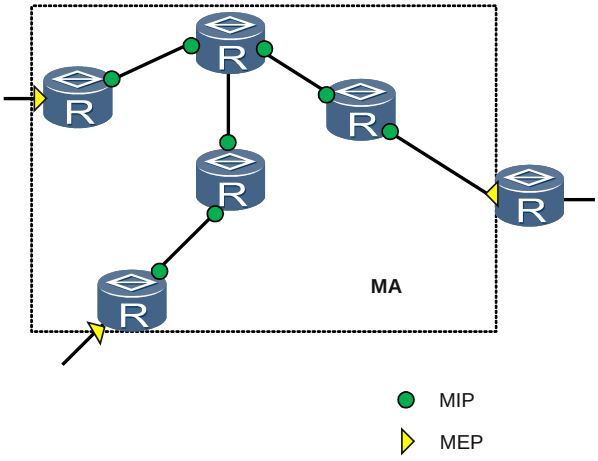
One or more MAs can be configured in an MD according to the actual needs. Each MA is a grouping of some maintenance points (MPs) in an MD. The level of an MA is the same as the level of the MD to which it belongs.

An MA serves a certain service (such as a VLAN): All packets sent by MPs in an MA carry the service tag, and an MP can receive packets sent by other MPs in the same MA.

MP

MPs are configured on ports and belong to a certain MA. There are two MPs: MEP and MIP. Figure 31-3 shows the MPs.

Figure 31-3 MP



MEP

MEPs are end points of the MD and MA, and they specify the range and boundary of the MD and MA.

The level of an MEP is identical to the level of the MD to which it belongs. The level of an MEP determines the level of packets it can process.

- When an MEP receives a packet whose level is higher, it does not process this packet and forwards the packet along its original path so that the packet can traverse the MD.
- If an MEP receives a packet whose level is the same or lower, the MEP will process the packet.

The levels of packets sent by an MEP are the same as the level of the MEP.

MEPs are located on ports of devices and need to be created by users manually.

The MEP on any device that runs Ethernet CFM is called the local MEP. The MEPs on the other devices in the same MA are the remote maintenance association end points (RMEPs) to the local MEP.

MIP

MIPs are inside an MA. Network management can be improved by deploying multiple MIPs among MEPs. More MIPs bring about stronger control and improved management over networks. For profitable services that key accounts care about, carriers deploy much more MIPs.

MIPs are located on ports of devices, and they are generated automatically according to specific rules but cannot be created manually.

31.4.4 CFM Principles

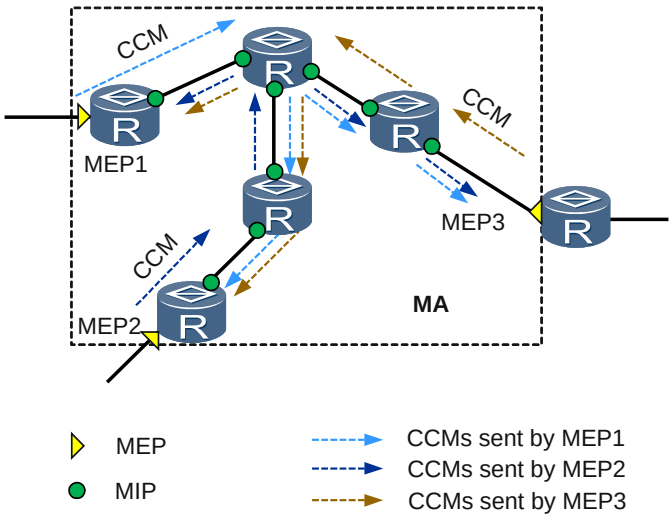
This topic describes the basic concepts and principles of CFM based on 802.1ag. Because the connectivity fault management (CFM) function defined by Y.1731 is similar to the one defined by 802.1ag, you can refer to the similar concepts of 802.1ag to get to know the mechanism of Y.1731 unless otherwise specified.

CC Principle

Continuity check (CC) is used to monitor connectivity of Ethernet links in real time and to report link exceptions by alarms.

A maintenance association end point (MEP) periodically sends multicast continuity check messages (CCMs) to other MEPs in the same maintenance association (MA). If an MEP fails to receive expected CCMs or receives unexpected CCMs within a certain period, the MEP considers that it has detected a connectivity fault in its MA. Figure 31-4 illustrates the CC principle.

Figure 31-4 CC principle



1. MEP database setup

   Each device with Ethernet connectivity fault management (CFM) enabled has an MEP database. The database records the MEPs (local MEPs) configured on the local device and other devices' MEPs (that is, RMEPs) in the same MA. Local MEPs and RMEPs are manually configured but automatically recorded to the MEP database.

2. CCM generation

   An MEP generates and sends CCMs. As shown in Figure 31-4, MEP1, MEP2, and MEP3 are in the same MA. After the function of sending CCMs is enabled, MEP1, MEP2, and MEP3 periodically multicast CCMs to each other at the same interval.

   All maintenance domain intermediate points (MIPs) and MEPs in an MA receive CCMs, with no need to send a reply.

   A CCM carries its maintenance level information. The level of a CCM is equal to the level of the MEP sending the CCM.

3. CCM termination

   An MEP terminates CCMs. If an MEP receives a CCM whose level is higher than the MEP's, the MEP forwards this CCM. If an MEP receives a CCM whose level is lower than or equal to the MEP's, the MEP does not forward this CCM, ensuring that CCMs in a lower-level maintenance domain (MD) are not sent to a higher-level MD.

4. Fault identification

   If an MEP fails to receive expected CCMs or receives unexpected CCMs within a certain period, the MEP considers that it has detected a connectivity fault in its MA and reports a

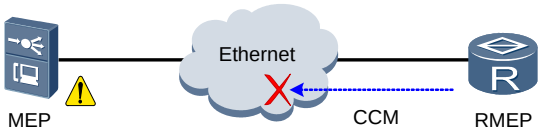
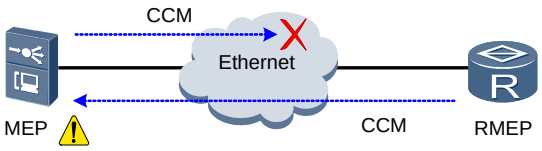
link fault alarm. When multiple MEPs in an MD send CCMs, multipoint-to-multipoint (MP2MP) link connectivity is checked.

NOTE

CC can determine whether a fault occurs but cannot identify the fault location. In this case, loopback (LB) and linktrace (LT) can be used to determine the fault location.

Table 31-4 describes the alarms may be generated when link faults occur on a network.

Table 31-4 CC Alarm

Alarm Name	Principle	Troubleshooting
0x2f210001 Loss of Ethernet connectivity check message	If an MEP does not receive a CCM from an RMEP within a period equal to 3.5 times the CC interval, the MEP considers that the RMEP is lost, and the device on which the MEP is configured generates an alarm. The following figure shows how the alarm is generated. 	When the alarm is generated, the link from the RMEP to the MEP is faulty. In this case, LB or LT can be used on the MEP to determine the fault location.
0x2f210004 Reception of Ethernet connectivity check message with the RDI bit set	When an MEP receives CCMs from its RMEP but the RMEP fails to receive CCMs from the MEP within a period equal to 3.5 times the CC interval, the RMEP considers that the MEP is lost and sets the RDI bit in its CCMs to be sent to 1. If the MEP receives a CCM with the RDI bit set to 1, the device on which the MEP is configured generates an alarm. The following figure shows how the alarm is generated. 	When the alarm is generated, the link from the MEP to the RMEP is faulty but the link from the RMEP to the MEP is normal. That is, a unidirectional connectivity fault occurs on the link. In this case, LB or LT can be used on the MEP to determine the fault location.
0x2f210002 Reception of invalid Ethernet connectivity check message	When an MEP receives CCMs with errors from its RMEP, the device on which the MEP is configured generates an alarm indicating reception of incorrect Ethernet OAM CCMs. Error types include: • Unexpected MEG level defect (dUNL): The level of CCMs is lower than the level of the MEP that receives the CCMs. • Unexpected Periodicity defect (dUNP): The interval of sending CCMs is different from the CC interval of the MEP that receives the CCMs. • Unexpected MEP defect (dUNM): The MEP ID of CCMs is not included in the RMEP ID list of the MEP that receives the CCMs.	When the alarm is generated, the link between the MEP and the RMEP is not interrupted but the MEP receives incorrect CCMs. In this case, the network CFM data plan needs to be checked.

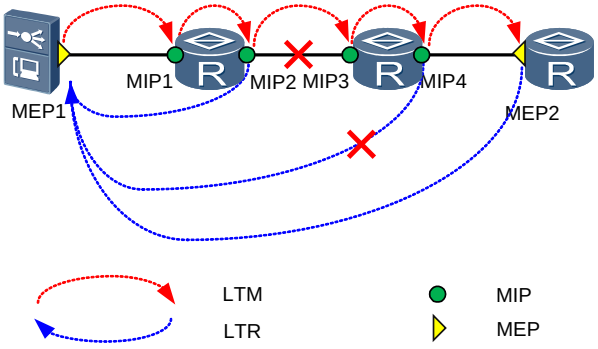
Alarm Name	Principle	Troubleshooting
0x2f210003 Reception of Ethernet cross connect connectivity check message	When a CCM received by an MEP from its RMEP contains an item (MD name type, MD name, MA name type, or MA name) different from that configured on the MEP, the device on which the MEP is configured generates an alarm indicating reception of cross connect Ethernet OAM CCMs.	When the alarm is generated, the link between the MEP and the RMEP is not interrupted but the MEP receives incorrect CCMs. In this case, the CFM data plan needs to be checked on the MEP and the RMEP.

LT Principle

Linktrace (LT) is used to obtain MAC addresses of intermediate devices along the Ethernet link between two devices and also determine the location of a link fault.

Along the link between the maintenance association end point (MEP) LT initiator and the target MEP, every device that generates a maintenance domain intermediate point (MIP) sends a linktrace reply (LTR) to the initiator and forwards the linktrace message (LTM) until the LTM reaches its target. The MEP LT initiator, through LTRs, obtains MAC addresses and locations of all in-between devices as well as the link section where a fault occurs. Figure 31-5 illustrates the LT principle.

Figure 31-5 LT principle



1. When all links are functioning properly, MEP1 sends an LTM to MEP2.
2. After all devices along the link between MEP1 and MEP2 receive the LTM, egress points (MIP2 and MIP4) return an LTR to MEP1, and forward the LTM with the time to live (TTL) field decremented by 1 to the next hop. An LTR contains the MAC address of the device sending the LTR.
3. After receiving the LTM, MEP2 stops forwarding the LTM but directly returns an LTR to MEP1.

4. As shown in Figure 31-5, when the link between MIP2 and MIP3 is faulty, MEP1 can receive an LTR from MIP2 but not from MIP4 after sending an LTM destined for MEP2. In this way, the fault location is determined.

LB Principle

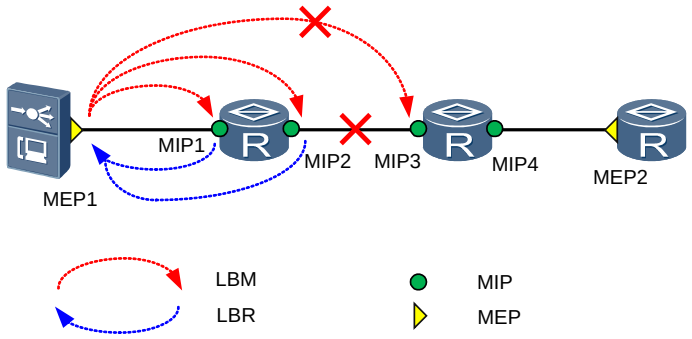
Loopback (LB) is used to check connectivity of end-to-end (E2E) Ethernet links between devices and also determine the location of a link fault.

Unicast LB

Unicast LB, similar to the ping operation at the IP layer, is mainly used to check the status of connection between a local device and a remote device.

Figure 31-6 illustrates the principle of unicast LB.

Figure 31-6 Unicast LB principle



1. MEP1 sends a loopback message (LBM) to MIP1.
2. MEP1 receives a loopback reply (LBR) from MIP1.

NOTE

A maintenance domain intermediate point (MIP) only responds to the received LBM but does not forward the LBM to the next-hop MIP or the target maintenance association end point (MEP).

3. MEP1 sends an LBM to the next hop of MIP1, that is, MIP2.
4. MEP1 receives an LBR from MIP2.
5. MEP1 sends an LBM to the next hop of MIP2, that is, MIP3.
6. MEP1 fails to receive an LBR from MIP3 because the link between MIP2 and MIP3 is faulty.
7. MEP1 determines that the link between MIP2 and MIP3 is faulty.

In unicast LB, the destination MAC address of an LBM is a unicast MAC address. Therefore, before using unicast LB, ensure that:

- The MAC address or MEP index of MEP2 and the MAC addresses of all MIPs along the link between MEP1 and MEP2 are obtained if the connection status of all links between MEP1 and MEP2 is required for accurately determining the fault location. LT can be performed on MEP1 to obtain the MAC addresses of all MIPs along the link between MEP1 and MEP2.
- The MAC address or MEP index of MEP2 is obtained if only the connection status between MEP1 and MEP2 is required.

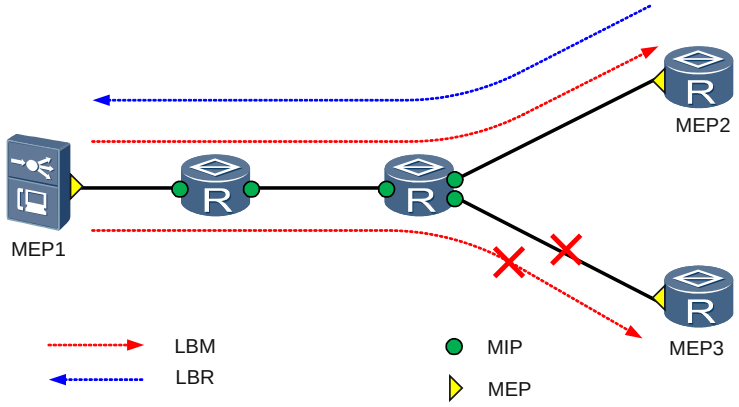
MEP1 can send multiple LBMs and, based on the number of received LBRs, determine whether packet loss occurs on the network.

Multicast LB

Multicast LB is used to check connectivity between an MEP and its peer MEPs and to obtain MAC addresses of those MEPs.

Multicast LB generally applies to scenarios in which multiple MEPs are deployed. In multicast LB, the MAC address of an LBM is a multicast MAC address in format of 01-80-C2-00-00-3x, where x is the level of the MEP initiating multicast LB. When the level of a peer MEP is equal to the level of the received multicast LBM, the peer MEP returns an LBR to the MEP initiating multicast LB. Figure 31-7 illustrates the principle of multicast LB.

Figure 31-7 Multicast LB principle



1. MEP1 multicast an LBM.
2. MEP2 returns an LBR after receiving the LBM.
3. MEP3 fails to receive the LBM and therefore does not return an LBR because the link between MEP3 and MEP1 is interrupted.
4. MEP1, according to the received LBR, determines that the link between MEP1 and MEP2 is functioning properly but the link between MEP1 and MEP3 is faulty.

An LBR contains the MAC address of the remote MEP (RMEP) sending this LBR; therefore, multicast LB ca n also be used to obtain the MAC address of an RMEP.

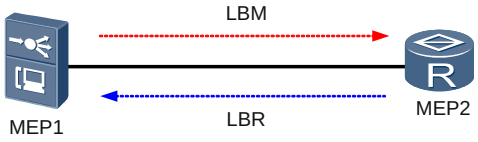
As defined in the Y.1731, only one multicast LBM can be sent at a time. However, for compatibility with LB defined in the 802.1ag, the MA5600T/MA5603T/MA5608T supports a configurable number of multicast LBMs to be sent. An RMEP has a delay in responding to the received LBM. Therefore, an MEP may receive disordered LBRs in response to multiple LBMs it sends. That is, the serial numbers (SNs) of LBRs returned by an RMEP may not be consecutive. Therefore, it is recommended that the number of LBMs that can be sent at a time be set to 1 before multicast LB is performed.

Bidirectional diagnostics test

Bidirectional diagnostics test is used to check the packet loss rate and bit error rate of a bandwidth-specific bidirectional link between two MEPs to determine whether the bandwidth meets user requirements.

Figure 31-8 illustrates the principle of bidirectional diagnostics test.

Figure 31-8 Bidirectional diagnostics test principle



In bidirectional diagnostics test, an LBM and an LBR both carry a testing code, which is used for pseudo random binary sequence (PRBS) check. The following describes the process of bidirectional diagnostics test:

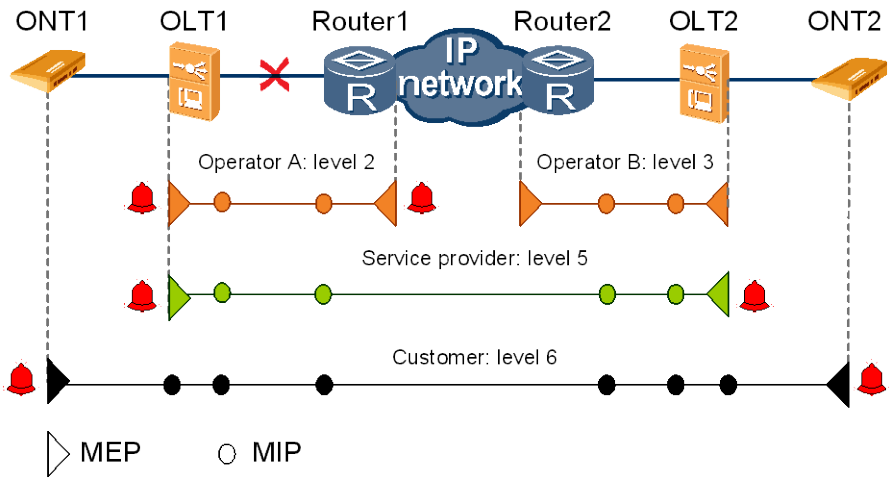
1. MEP1 sends a unicast LBM to MEP2.
2. MEP2 modifies the received LBM to an LBR and returns it to MEP1.
3. MEP1 analyzes the received LBR and determines whether packet loss and bit errors occur over the bidirectional link between MEP1 and MEP2.

AIS Principles

Alarm indication signal (AIS) is a dedicated function of Y.1731 and is not supported by 802.1ag. AIS is used to send link fault notifications to the client. AIS also suppresses redundant alarms.

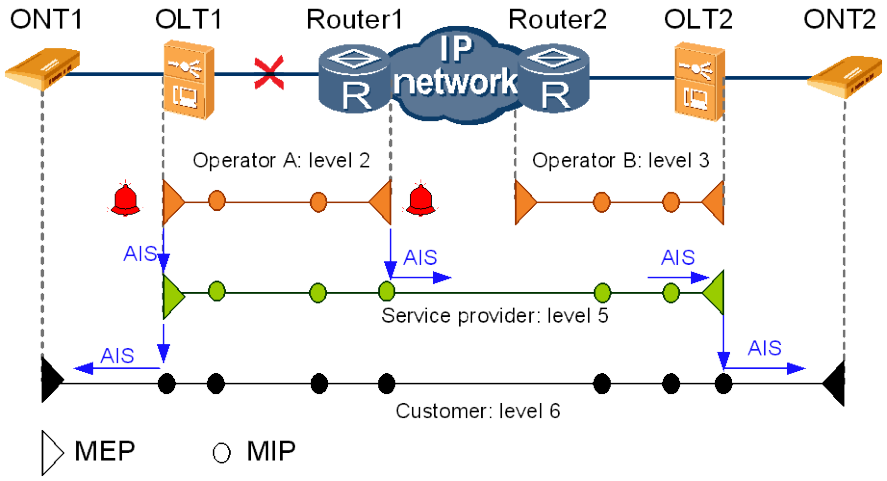
As shown in Figure 31-9, when lower-level services encounter continuity check (CC) faults, maintenance association end points (MEPs) at each level detect CC faults through ETH CC and report fault alarms. Alarms generated in this scenario are mostly redundant alarms. Actually, the OAM engineer can locate the section of the link where the fault has occurred according to the alarm reported by the lowest-level MEP. For example, if the network of Operator A encounters a fault, the fault can be identified within this network, and there is no need to locate the fault in maintenance domains (MDs) owned by the service provider and customer.

Figure 31-9 Before AIS Is Enabled



Alarm suppression can be performed through AIS. Specifically, after a lower-level MEP detects a CC fault, this MEP periodically sends AIS packets to the client MEG level in the reverse direction. After receiving these AIS packets, the client-level MEP no longer reports an alarm when detecting a CC fault. For example, MDs owned by the service provider and customer do not report any alarms.

Figure 31-10 After AIS Is Enabled



Fault types triggering AIS alarms include connection loss, crossed connection, errored MEP ID, errored period, and MAC status fault.

LLF Principles

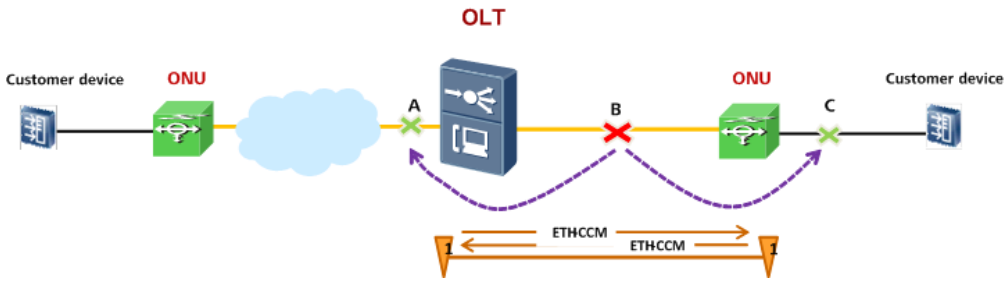
Link loss forwarding (LLF) is a Huawei-defined Ethernet OAM function. It transmits fault information about Ethernet links according to message types defined in the 802.1ag protocol.

Usually, LLF is used in open access networks for national broadband programs. In fiber to the office (FTTO) scenarios, retail service providers (RSPs) set up networks of their own and rent access pipelines and device ports from operators. If pipelines rent to RSPs encounter faults, LLF can be used to notify RSPs and the RSPs' affected enterprise users so that they can know about the faults and take corresponding measures more responsively.

LLF sends the interface status TLV (defined in 802.1ag) carried in the continuity check messages (CCMs) to the peer maintenance end point (MEP). Then, the peer MEP judges whether to generate a fault alarm or fault clear alarm according to the received interface status TLV.

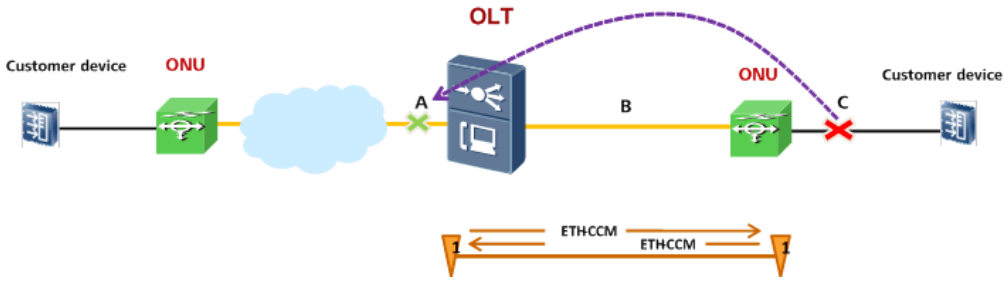
1. Maintenance scenario 1: The network is disconnected and the network-side port and user-side port are disabled. The network continuity is checked using the continuity check (CC) function of the MEP. Specifically, if the MEP on port A or port C detects a CC failure, port A or port C will be disabled; if the MEP on port A or port C detects a CC failure recovery, port A or port C will be enabled. Figure 31-11 shows this scenario.

Figure 31-11 Disabling the network-side port and user-side port when the network is disconnected



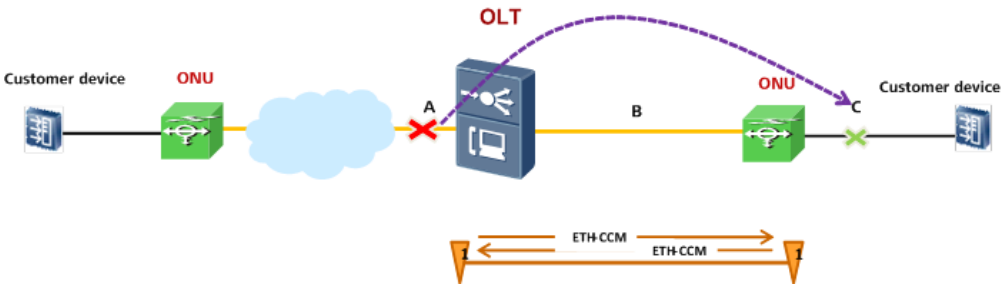
2. Maintenance scenario 2: The user-side port is disconnected and the corresponding network-side port is disabled. Specifically, when the user-side port C is disconnected, **interface status TLV** carried in the CCM is used to indicate whether the local port is faulty. If the MEP on port A receives a CCM carrying **interface status TLV 2**, the MEP will be notified that the user-side port C is faulty and will disable port A; if the MEP on port A receives a CCM carrying **interface status TLV 1**, the MEP will be notified that the user-side port C has recovered and will enable port A. Figure 31-12 shows this scenario.

Figure 31-12 Disabling the corresponding network-side port when the user-side port is disconnected



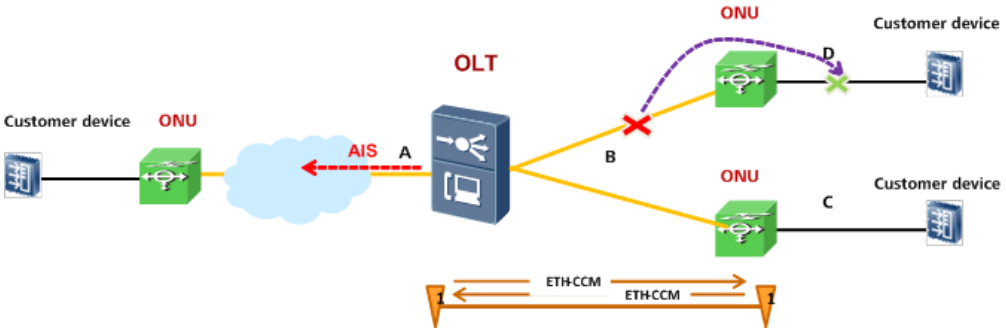
3. Maintenance scenario 3: The network-side port is disconnected and the corresponding user-side port is disabled. Specifically, when the network-side port A is disconnected, **interface status TLV** carried in the CCM is used to indicate whether the local port is faulty. If the MEP on port C receives a CCM carrying **interface status TLV 2**, the MEP will be notified that the network-side port A is faulty and will disable port C; if the MEP on port C receives a CCM carrying **interface status TLV 1**, the MEP will be notified that the network-side port A has recovered and will enable port C. Figure 31-13 shows this scenario.

Figure 31-13 Disabling the corresponding user-side port when the network-side port is disconnected



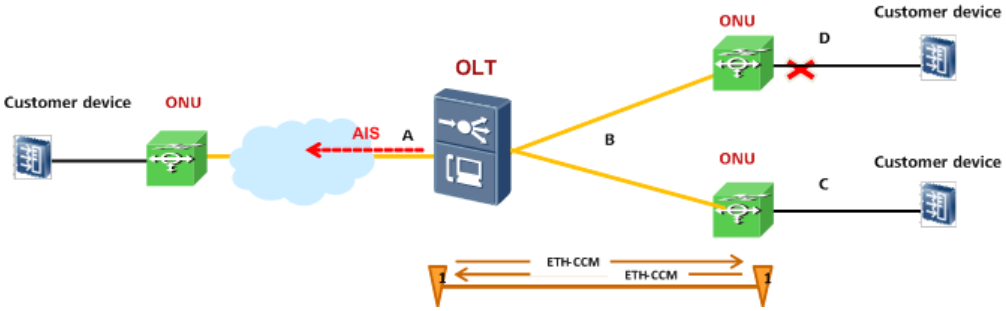
4. Maintenance scenario 4: In the aggregation scenario, the network is disconnected and the alarm indication signal (AIS) is sent. Specifically, in the aggregation scenario, when services of a user in the network are interrupted, the network-side port cannot be disabled because services of the other users are still normal. In this case, the OLT sends the AIS to the normal ONU to notify a link abnormality. If link B is disconnected, the MEP on port A detects a failure and notifies the ONU of this failure; if the MEP on port D detects a CC failure, port D is disabled. After the fault is rectified, if the MEP on port D detects a CC failure recovery, port D will be enabled; if the MEP on port A detects a CC failure recovery, it stops sending AISs. Figure 31-14 shows this scenario.

Figure 31-14 Sending AISs when the network is disconnected in the aggregation scenario



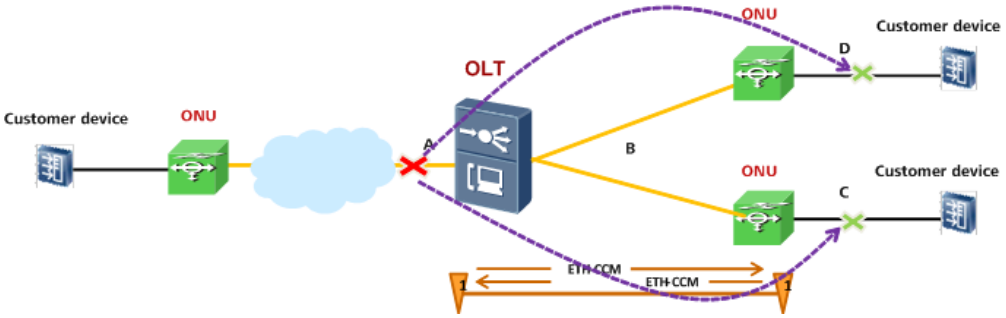
5. Maintenance scenario 5: In the aggregation scenario, the user-side port is disconnected and the AIS is sent. Specifically, in the aggregation scenario, when the user-side link is disconnected, the network-side port cannot be disabled because the other user-side links are still normal. In this case, the AIS is sent to notify the ONU of a link abnormality. When user-side port D is disconnected, **interface status TLV** carried in the CCM is used to indicate whether the local port is faulty. If the MEP on port A receives a CCM carrying **interface status TLV 2**, the MEP will be notified that the user-side port D is faulty and will send the AIS to notify the ONU of a failure; if the MEP on port A receives a CCM carrying **interface status TLV 1**, the MEP will be notified that the user-side port D has recovered and will stop sending the AIS. Figure 31-15 shows this scenario.

Figure 31-15 Sending AISs when the user-side port is disconnected in the aggregation scenario



6. Maintenance scenario 6: In the aggregation scenario, the network-side port is disconnected and all the user-side ports are disabled. Specifically, in the aggregation scenario, when the network-side port A is disconnected, **interface status TLV** carried in the CCM is used to indicate whether the local port is faulty. If the MEP on port C or port D receives a CCM carrying **interface status TLV 2**, the MEP will be notified that the port C or port D is faulty and will disable port C or port D; If the MEP on port C or port D receives a CCM carrying **interface status TLV 1**, the MEP will be notified that port C or port D has recovered and will enable port C or port D. Figure 31-16 shows this scenario.

Figure 31-16 Disabling all the user-side ports when the network-side port is disconnected in the aggregation scenario



The preceding scenarios which are achieved under the guideline of the following LLF principles:

1. The local MEP adds interface status TLV to the CCM according to the port link status and then sends the CCM.
2. The peer MEP determines to generate or not to generate a someMACstatusDefect alarm according to the interface status TLV carried in the CCM.

   The someMACstatusDefect alarm is generated by complying with the 802.1ag protocol. Specifically, when the port on a remote MEP is not UP, the MEP generates the someMACstatusDefect alarm. Alternatively, when none of the ports on the remote MEPs are in the forwarding state, the MEPs generate the someMACstatusDefect alarm.

Defect		Priority	
Variable	highestDefect (20.33.9)	highestDefectPri (20.33.8)	Importance
xconCCMdefect (20.23.3)	DefXconCCM	5	Most
errorCCMdefect (20.21.3)	DefErrorCCM	4	
someRMEPCCMdefect (20.33.5)	DefRemoteCCM	3	
someMACstatusDefect (20.33.6)	DefMACstatus	2	
someRDIdefect (20.33.7)	DefRDICCM	1	Least

3. When an MEP detects a CC alarm that has a higher severity than the associated alarm severity, the port connected to the MEP is automatically shut down. When the alarm is cleared, the port status recovers. A port can also be shut down through the CLI. Therefore, before performing configuration, know the relationships between the associated port shutdown function and the CLI-configured port shutdown function, as listed in the following table.

Associated Status	CLI-configured Status	Final Status
Shutdown	Shutdown	Shutdown
Shutdown	Undo shutdown	Shutdown
Undo shutdown	Shutdown	Shutdown
Undo shutdown	Undo shutdown	Undo shutdown
Associated port shutdown function disabled	Shutdown	Shutdown
Associated port shutdown function disabled	Undo shutdown	Undo shutdown

NOTE

Triggering the associated port shutdown function means to associate the MEP CC alarm status with the activation status of the port where the MEP resides. In other words, the port is automatically disabled or enabled according to the CC status. When association is enabled, the lowest CC alarm severity (2 by default) needs to be configured. Association is triggered when the severity of the CC alarm generated by the MEP is higher than or equal to the configured association alarm severity. When association is disabled, the port status is irrelevant to the CC status.

31.4.5 Configuring the Ethernet CFM OAM

CFM OAM is an OAM at the network level. It supports connectivity check, loopback testing and link tracking, and applies to the end-to-end fault detection in large-scale network.

Prerequisites

- ONT or Modem must support the 802.1ag protocol.
- Service configurations (VLAN configuration and service port configuration, for example) are finished, and the customer services are normal.

Context

The following table shows basic concepts for CFM OAM.

Concepts	Explanations
Maintenance Domain (MD)	Maintenance domain (MD) is a network or part of a network where the CFM management is implemented. The MD can be divided into three types according to the size of the range: customer domain, service provider domain, and operator domain.
Maintenance Association (MA) service	Maintenance association (MA) is a part of an MD, a instance of an MD, and it associates with the monitoring service. An MD consists of one or more MAs.
Maintenance association End Point (MEP)	MEP is the end point of an MA. For any device in the network running Ethernet CFM, the MEP on the device is called as local MEP, MEPs on other devices in the same MA are called as remote MEP (relative to the local MEP).
MD Level	There are eight levels, from 0 to 7. It is carried in CFM packets. CFM packets with high-level can pass through the low-level MD. Therefore, different levels of MDs can be nested deployed.

For details about the principle of the CFM OAM, please refer to 31 Ethernet OAM.

Networking

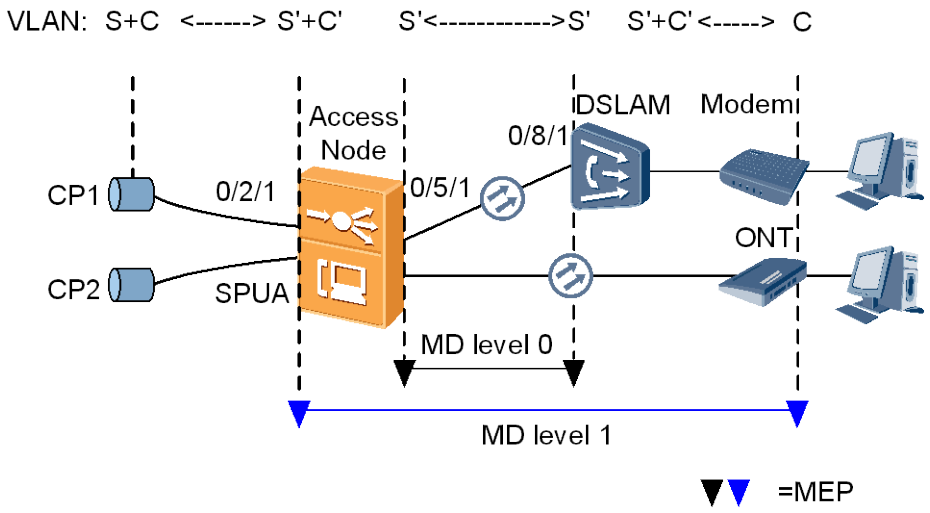
Figure 31-17 shows a typical FTTx integrated networking. Where,

- MA5600T/MA5603T/MA5608T, functioning as access node, accesses the cascading DSLAM by SPUA board.
- MA5600T/MA5603T/MA5608T also functions as OLT, and provides multi-services for customers by accessing ONT.
- DSLAM provides multi-services for customers by accessing xDSL modem.

Through the deployment of CFM OAM, service providers can detect connectivity between any two devices based on their needs. When there is a connectivity problem, the system generates an alarm reporting the fault location.

- Deploy the MD for management channel between cascading port on MA5600T/MA5603T/MA5608T and upstream port on DSLAM.
- Deploy another MD for service channel between upstream port on MA5600T/MA5603T/MA5608T and UNI port of ONT (or Modem). The level of this MD is higher than the former.

Figure 31-17 Typical networking of CFM OAM



Data Plan

Table 31-5 shows the key data plan for deploying the MD for management channel between cascading port on MA5600T/MA5603T/MA5608T and upstream port on DSLAM.

Table 31-5 Data plan for the Ethernet CFM OAM - MD for management channel

Item	Data
Access Node: MA5600T/MA5603T/MA5608T	• MD ID: 0. MD name: fttc_md0 • MD level: 0 • MA ID: 0. MA name: fttc_ma0 • MA VLAN: 100 (Management VLAN for the MA5600T/MA5603T/MA5608T) • MEP ID: 1. MEP port: 0/2/1 • MEP VLAN tag: 8 (Management VLAN for the DSLAM). MEP direction: Down • Remote MEP ID: 2 • CC-interval: 10 minutes
DSLAM (supposed to be an MA5600T/MA5603T/MA5608T)	• MD ID: 0. MD name: fttc_md0 • MD level: 0 • MA ID: 0. MA name: fttc_ma0 • MA VLAN: 8 (Management VLAN for the DSLAM) • MEP ID: 2. MEP port: 0/3/1 • MEP direction: Down • Remote MEP ID: 1 • CC-interval: 10 minutes

Table 31-6 shows the key data plan for deploying another MD for service channel between upstream port on MA5600T/MA5603T/MA5608T and UNI port of ONT (or Modem).

Table 31-6 Data plan for the Ethernet CFM OAM - MD for service channel

Item	Data
Access Node: MA5600T/MA5603T/MA5608T	• MD ID: 1. MD name: fttc_md1 • MD level: 1 • MA ID: 1. MA name: fttc_ma1 • MA VLAN: 0 (indicates the MA does not associate any VLAN in system) • MEP ID: 1. MEP port: 0/2/1 • MEP VLAN tag1: 1000 (SVLAN for data service on the upstream port of the SPUA board) • MEP VLAN tag1: 10 (Inner VLAN for data service on the upstream port of the SPUA board) • MEP direction: Up • Remote MEP ID: 2 • CC-interval: 10 minutes
UNI interface of the ONT (or Modem)	Fixed MEP with its level 1.

Procedure

- Configure the MD for management channel on the MA5600T/MA5603T/MA5608T (MD index 0 and MD level 0).

  Pay attention to the follows:

  - The MD name and MA name configured on MA5600T/MA5603T/MA5608T must be the same as that of on DSLAM.
  - Local MEP on MA5600T/MA5603T/MA5608T corresponds to the remote MEP on the DSLAM, and remote MEP on MA5600T/MA5603T/MA5608T corresponds to the local MEP on the DSLAM.

  a. Configure the MD.

  Configure MD 0 with a name of the character string type, name fttc_md0, and MD level 0.

  - MDs with the same index or level cannot be created.
  - The name type and the name of an MD must be unique.
  - The total length of the names of an MD and its MAs cannot be longer than 44 characters.
  - The MD name type, the MD name and the MD level must be consistent at both ends.

```
huawei(config)#cfm md 0 name-format string fttc_md0 level 0 mhf-creation no-mhf
```

  b. Configure the MA.

- The system supports up to 4096 MAs. That is, if an MD is configured with 4096 MAs, the other MDs in the system cannot be configured with any MA.
- An MD of must be available for creating an MA.
- An existing MA cannot be created again.
- The total length of the names of an MD and its MAs cannot be longer than 44 characters.
- The MA name type, the MA name and the sending period of CC packets must be consistent at both ends.

Create an MA with the index 0/0. The name type is the character string type, and the name is fttc_ma0. The sending period of CC packets is 10 minutes (the sending period of CC packets is 1 minute by default).

```
huawei(config)#cfm ma 0/0 name-format string fttc_ma0 cc-interval 10m
```

Set the VLAN associated to the MA to 100, it is the management VLAN of the MA5600T/MA5603T/MA5608T.

```
huawei(config)#cfm ma 0/0 vlan 100
```

Set the ID of MEP contained by the MA to 1 and 2. Currently, an MA supports a local MEP and a remote MEP, and their IDs must be unique. MEP ID 2 needs to be configured on the peer DSLAM.

```
huawei(config)#cfm ma 0/0 meplist 1  //local end MEP
huawei(config)#cfm ma 0/0 meplist 2 //remote end MEP on the DSLAM
```

c. Configure the MEP.

- MEP refers to a maintenance association end points. Ethernet CFM OAM is used to test the link connectivity by using the MEPs at the two ends of a maintenance channel.
- By default, the MEP management function is enabled, the priority of sending CFM packets is 7, and the function of sending CC packets is enabled.
- There are two kinds of MEPs: UP MEP and DOWN MEP. An UP MEP indicates that the MEP transmits packets to the bridge trunk direction. A DOWN MEP indicates that the MEP transmits packets to the physical medium direction.
- **vlantag1** or **vlantag2** must be configured, when you add an MEP is added for a port with service streams. **vlantag1** is the outer VLAN of the port carrying the service link for the MEP. **vlantag2** is the inner VLAN of the port carrying the service link for the MEP.
- The MEP priority must be consistent at both ends.

```
huawei(config)#cfm mep 0/0/1 direction down port 0/2/1 vlantag1 8 priority 7
```

d. Enable the remote MEP detection function.

The system can check the remote MEPs of an MA and report alarms for loss of CCM and RDI only when the following functions are enabled: the global CFM function, the global function of checking remote MEPs, and the function of checking the remote MEPs of the MA.

By default, the remote MEP detection function of MA is enabled, while the global remote MEP detection function is disabled.

i. Enable the remote MEP detection function of the MA.

```
huawei(config)#cfm ma 0/0 remote-mep-detect enable
```

ii. Enable the continuity check function of the MEP.

```
huawei(config)#cfm mep 0/0/1 cc enable
```

iii. Enable the global remote MEP detection function.

```
huawei(config)#cfm remote-mep-detect enable
```

e. Enable the global CFM function.

```
huawei(config)#cfm enable
```

- Configure the MD for management channel on the DSLAM (MD index 0 and MD level 0).

a. Configure the MD.

```
huawei(config)#cfm md 0 name-format string fttc_md0 level 0 mhf-creation no-mhf
```

b. Configure the MA.

```
huawei(config)#cfm ma 0/0 name-format string fttc_ma0 cc-interval 10m
```

Set the VLAN associated to the MA to 8, it is the management VLAN of the DSLAM.

```
huawei(config)#cfm ma 0/0 vlan 8
```

Set the ID of MEP contained by the MA to 2 and 1. Currently, an MA supports a local MEP and a remote MEP, and their IDs must be unique. MEP ID 1 needs to be configured on the peer MA5600T/MA5603T/MA5608T.

```
huawei(config)#cfm ma 0/0 meplist 2  //local end MEP on the DSLAM
huawei(config)#cfm ma 0/0 meplist 1 // remote end MEP on the MA5600T/MA5603T/MA5608T
```

c. Configure the MEP.

```
huawei(config)#cfm mep 0/0/2 direction down port 0/3/1 vlantag1 8 priority 7
```

d. Enable the remote MEP detection function.

i. Enable the remote MEP detection function of the MA.

```
huawei(config)#cfm ma 0/0 remote-mep-detect enable
```

ii. Enable the continuity check function of the MEP.

```
huawei(config)#cfm mep 0/0/2 cc enable
```

iii. Enable the global remote MEP detection function.

```
huawei(config)#cfm remote-mep-detect enable
```

e. Enable the global CFM function.

```
huawei(config)#cfm enable
```

- Configure the MD for service channel on the MA5600T/MA5603T/MA5608T (MD index 1 and MD level 1).

a. Configure the MD.

Configure MD 1 with a name of the character string type, name fttc_md1, and MD level 1.

```
huawei(config)#cfm md 1 name-format string fttc_md1 level 1 mhf-creation no-mhf
```

b. Configure the MA.

Create an MA with the index 1/1. The name type is the character string type, and the name is fttc_ma1. The sending period of CC packets is 10 minutes (the sending period of CC packets is 1 minute by default).

```
huawei(config)#cfm ma 1/1 name-format string fttc_ma1 cc-interval 10m
```

This level of ETH OAM is based on point to point forwarding service, the MA VLAN should be 0, it means that the MA does not associate any VLAN in the system.

```
huawei(config)#cfm ma 1/1 vlan 0
```

Set the ID of MEP contained by the MA to 1 and 2. Currently, an MA supports a local MEP and a remote MEP, and their IDs must be unique. MEP ID 2 needs to be configured on the peer device.

```
huawei(config)#cfm ma 1/1 meplist 1   //local end MEP
huawei(config)#cfm ma 1/1 meplist 2   //remote end MEP on the ONT (or Modem)
```

c. Configure the MEP.

Set the service VLAN to 1000 and inner vlan to 10 on the upstream port of SPUA board, which maps to an UNI port of the ONT (or Modem). Set the direction to up.

```
huawei(config)#cfm mep 1/1/1 direction up port 0/2/1 vlantag1 1000 vlantag2
10 priority 7
```

d. Enable the remote MEP detection function.

The system can check the remote MEPs of an MA and report alarms for loss of CCM and RDI only when the following functions are enabled: the global CFM function, the global function of checking remote MEPs, and the function of checking the remote MEPs of the MA.

By default, the remote MEP detection function of MA is enabled, while the global remote MEP detection function is disabled.

i. Enable the remote MEP detection function of the MA.

```
huawei(config)#cfm ma 1/1 remote-mep-detect enable
```

ii. Enable the continuity check function of the MEP.

```
huawei(config)#cfm mep 1/1/1 cc enable
```

iii. Enable the global remote MEP detection function.

```
huawei(config)#cfm remote-mep-detect enable
```

e. Enable the global CFM function.

```
huawei(config)#cfm enable
```

- Configure the alarm indication signal (AIS) function.

a. Set the client level of MA 0/0 to 1.

```
huawei(config)#cfm ma 0/0 client-level 1
```

b. Enable AIS transmission on MEP 0/0/1.

```
huawei(config)#cfm mep 0/0/1 ais enable
```

c. Set the interval to 1 minute for sending AIS packets from MA 0/0.

```
huawei(config)#cfm ma 0/0 ais-interval 1m
```

d. Run the **display cfm ma** command to query configurations.

- Configure the link loss forwarding (LLF) function.

Set CC alarm severity to 3 at which port status association of MEP 0/0/1 is triggered.

```
huawei(config)#cfm mep 0/0/1 trigger if-down 3
```

----End

Result

After the configuration is finished,

- MA5600T/MA5603T/MA5608T or DSLAM is able to learn the MEP ID and MAC address from its remote peer automatically. You can run the **display cfm mep** command to query MEP configuration.
- Disconnect MA5600T/MA5603T/MA5608T and DSLAM, the system will generate CFM OAM alarm automatically, reporting the fault location and cause.
- Run the **cfm loopback** command on the MA5600T/MA5603T/MA5608T to start a remote loopback test. Under normal circumstances, the number of packets sent and received must be the same.

Configuration File

Configure the MD for management channel on the MA5600T/MA5603T/MA5608T (MD index 0 and MD level 0).

```
cfm md 0 name-format string fttc_md0 level 0 mhf-creation no-mhf
cfm ma 0/0 name-format string fttc_ma0 cc-interval 10m
cfm ma 0/0 vlan 100
cfm ma 0/0 meplist 1   //local end MEP
cfm ma 0/0 meplist 2 //remote end MEP on the DSLAM
cfm mep 0/0/1 direction down port 0/2/1 vlantag1 8 priority 7
cfm ma 0/0 remote-mep-detect enable
cfm mep 0/0/1 cc enable
cfm remote-mep-detect enable
cfm enable
```

Configure the MD for management channel on the DSLAM (MD index 0 and MD level 0).

```
cfm md 0 name-format string fttc_md0 level 0 mhf-creation no-mhf
cfm ma 0/0 name-format string fttc_ma0 cc-interval 10m
cfm ma 0/0 vlan 8
cfm ma 0/0 meplist 2   //local end MEP
cfm ma 0/0 meplist 1 //remote end MEP on the MA5600T/MA5603T/MA5608T
cfm mep 0/0/2 direction down port 0/3/1 vlantag1 8 priority 7
cfm ma 0/0 remote-mep-detect enable
cfm mep 0/0/2 cc enable
cfm remote-mep-detect enable
cfm enable
```

Configure the MD for service channel on the MA5600T/MA5603T/MA5608T (MD index 1 and MD level 1).

```
cfm md 1 name-format string fttc_md1 level 1 mhf-creation no-mhf
cfm ma 1/1 name-format string fttc_ma1 cc-interval 10m
cfm ma 1/1 vlan 0
cfm ma 1/1 meplist 1   //local end MEP
cfm ma 1/1 meplist 2 //remote end MEP on the DSLAM
cfm mep 1/1/1 direction up port 0/2/1 vlantag1 1000 vlantag2 10 priority 7
cfm ma 1/1 remote-mep-detect enable
```

```
cfm mep 1/1/1 cc enable
cfm remote-mep-detect enable
cfm enable
```

Configure AIS transmission on MA 0/0.

```
cfm ma 0/0 client-level 1
cfm mep 0/0/1 ais enable
cfm ma 0/0 ais-interval 1m
```

Configure the LLF function, and set CC alarm severity to 3 at which port status association of MEP 0/0/1 is triggered.

```
cfm mep 0/0/1 trigger if-down 3
```

31.5 EFM (802.3ah)

EFM is defined in the IEEE 802.3ah (802.3ah for short). It is an OAM feature for detecting quality and connectivity of last-mile Ethernet links.

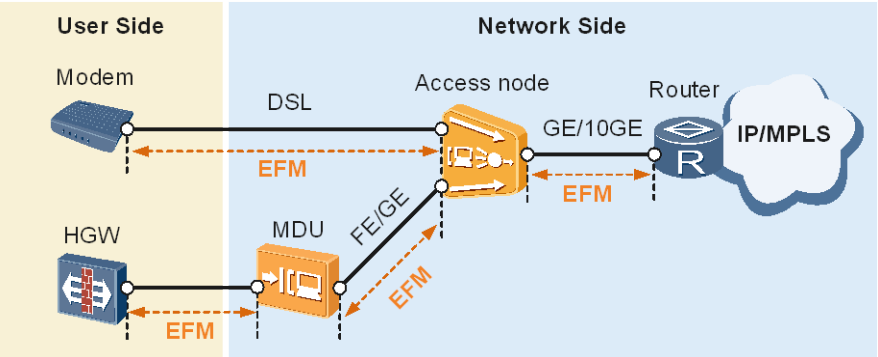
31.5.1 EFM Introduction

As an OAM mechanism, EFM monitors the link status, quickly locates a faulty link, and identifies the fault type in real time for carriers.

Among Ethernet applications, high-end users, such as banks and financial companies, demand high reliability. They expect carriers to monitor both carrier networks and last-mile links that connect users to those carrier networks. However, this is a difficulty for carriers because user devices are out of control of carriers.

Ethernet in the First Mile (EFM) can be used to satisfy these demands. As an OAM mechanism, EFM monitors the link status, quickly locates a faulty link, and identifies the fault type in real time for carriers. EFM mainly applies to the user access network and also to the Ethernet physical link that directly connects two devices. The following example illustrates EFM implementation on the network shown in Figure 31-18.

Figure 31-18 Typical EFM network



The router and access node are placed in the CO, MDU is in the corridor or curb, modem and home gateway (HGW) are in users' home. By deploying EFM between the devices shown in Figure 31-18, network maintenance engineers can remotely check connectivity and quality of links between these devices.

Table 31-7 lists the main functions of EFM supported by access device.

Table 31-7 Main functions of EFM

Function	Purpose
OAM discovery	Negotiates EFM capabilities of the connected two ends and obtains the EFM status of the remote end.
Remote failure indication	Sends a remote failure indication to the remote end when traffic is interrupted because the device is faulty or is invalid to monitor the link connectivity.
Remote loopback	Loops back the packets sent to the remote end to the local end and tests the link by comparing the received and transmitted packets to obtain the link performance data (including packet loss rate).
Link monitoring	Monitor the link quality by sending a link deterioration event to the network maintenance engineer.

31.5.2 EFM Basic Concept

The following concepts are used a lot in the 31.5.3 EFM Principle. If you are not familiar with EFM, it is recommended to have a rough understanding of these basic concepts.

DTE

The data terminating entity (DTE) is the basic object on which EFM feature takes effect. A DTE is associated with a port.

- If the port is on the local device, it is a local DTE.
- If the port is on the remote device, it is a remote DTE.

OAMPDUs

EFM works at the data link layer and uses protocol packets called OAM Protocol Data Units (PDUs). EFM DTEs periodically exchange OAMPDUs to report link status, helping network administrators effectively manage networks. Figure 31-19 shows the OAMPDU format and common types of OAMPDUs. Table 31-8 lists and describes fields in an OAMPDU.

Figure 31-19 OAMPDU format

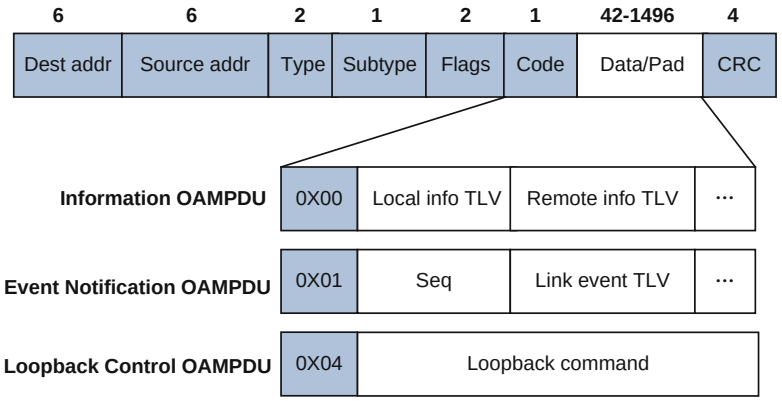


Table 31-8 Fields and descriptions in an OAMPDU

Field	Description
Dest addr	Destination MAC address, which is a slow-protocol multicast address 0x0180-C200-0002. Network bridges cannot forward slow-protocol packets thus EFM OAMPDUs cannot be forwarded over multiple devices.
Source addr	Source address, which is a unicast MAC address of a port on the transmit end. If no port MAC address is specified on the transmit end, the bridge MAC address of the transmit end is used.
Type	Slow protocol type, which has a fixed value of 0x8809.
Subtype	Subtype of a slow protocol. The value is 0x03, which means that the slow sub-protocol is EFM.
Flags	Status of a DTE: • Remote Stable • Remote Evaluating • Local Stable • Local Evaluating • Critical Event • Dying Gasp • Link Fault
Code	OAMPDU type: • 0X00: Information OAMPDU • 0X01: Event Notification OAMPDU • 0X04: Loopback Control OAMPDU Table 31-9 lists common types of OAMPDUs.

Table 31-9 OAMPDU types

OAMPDU Type	Description
Information OAMPDU	• Used to discover a remote DTE, initiate a handshake process, and establish an EFM connection. After the EFM connection is established, both EFM DTEs periodically exchange Information OAMPDUs to monitor link connectivity. • Used to advertise fault information. If the local DTE detects a fault, an event is generated on the local device and an Information OAMPDU is sent to the remote DTE. When the remote DTE receives the Information OAMPDU, an event is generated on the remote device.
Event Notification OAMPDU	Used to monitor links. If a local DTE detects an errored frame event, errored symbol period event, or errored frame second summary event, an event is generated on the local device and an Event Notification OAMPDU is sent to the remote DTE. When the remote DTE receives the Event Notification OAMPDU, an event is generated on the remote device.
Loopback Control OAMPDU	Used to enable or disable the remote loopback function.

Modes

EFM supports two modes: active and passive. Table 31-10 lists capabilities for processing OAMPDUs in the two modes.

Table 31-10 Capabilities for processing OAMPDUs in active and passive modes

Capability	Active Mode	Passive Mode
Initiate a connection request by sending an Information OAMPDU during the discovery process	Supported	Not supported
Respond to a connection request during the discovery process	Supported	Supported
Send Information OAMPDUs	Supported	Supported
Send Event Notification OAMPDUs	Supported	Supported
Send Loopback Control OAMPDUs	Supported	Not supported
Respond to Loopback Control OAMPDUs	Supported (The remote DTE must work in active mode.)	Supported

NOTE

- Two DTEs on both ends of an EFM link work in active mode by default.

- An EFM connection can only be initiated by a DTE working in active mode. A DTE working in passive mode waits to receive a connection request from its remote DTE. Two DTEs both working in passive mode cannot establish an EFM connection between them.
- A loopback request can only be initiated by a DTE working in active mode.

31.5.3 EFM Principle

EFM functions are defined in section 57 of the IEEE 802.3ah protocol. This topic describes only functions supported by the access device.

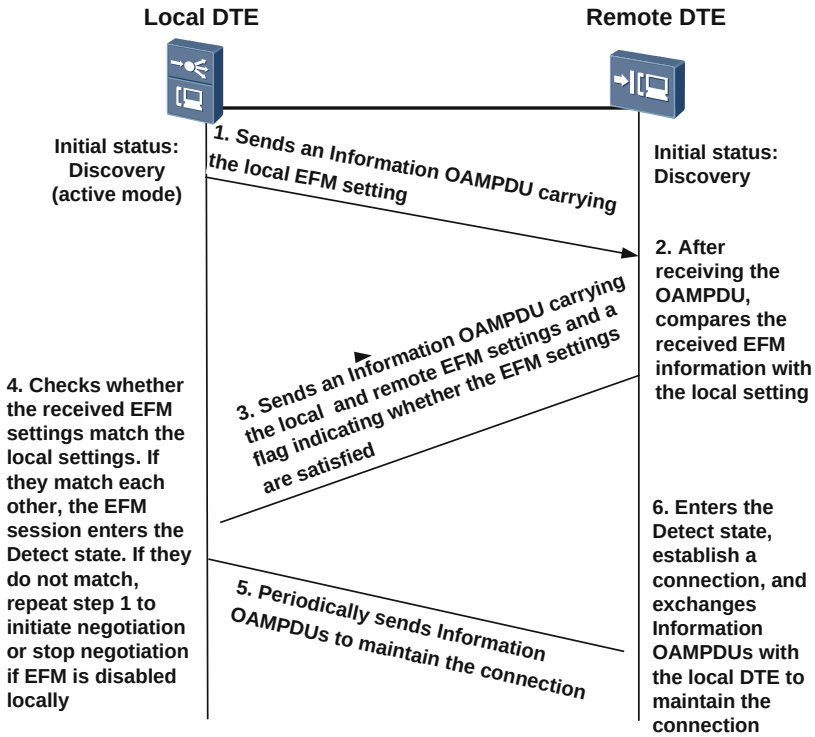
NOTE

The following principles are for the access device. Different devices support different functions. For details, see "Specifications."

OAM Discovery

After being enabled with EFM, the local DTE initiates the Discovery negotiation. During the Discovery phase, a local DTE discovers and establishes a stable EFM connection with a remote DTE. If both the local and remote DTEs are satisfactory for EFM functions, the EFM connection is established. Figure 31-20 shows this process.

Figure 31-20 Schematic diagram for OAM discovery



After the EFM connection is established, DTEs at both ends of an EFM connection periodically exchange Information OAMPDUs to monitor link connectivity. The interval at which Information OAMPDUs are sent is also known as an interval between handshakes. If a DTE does not receive Information OAMPDUs from the remote DTE within the connection timeout period, the DTE considers the connection interrupted and generates an event. Establishing an EFM connection is a way to monitor physical link connectivity automatically.

Remote Failure Indication

After the OAM discovery phase finishes, two DTEs at both ends of an EFM connection exchange Information OAMPDUs to monitor link connectivity. If traffic is interrupted due to a remote device failure, the remote DTE sends an Information OAMPDU carrying an event listed in Table 31-11 to the local DTE. After receiving the notification, the local DTE generates the Ethernet OAM link event. An administrator can determine link status based on the event and take measures to rectify the fault.

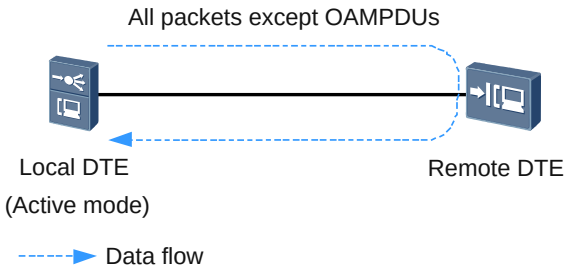
Table 31-11 Critical link events

Type	Description
Link Fault	A loss of signal (LoS) error occurs because the interval at which OAMPDUs are sent elapses or a physical link fails.
Dying Gasp	An unexpected status changes or event occurs because a remote device is powered off or a board is reset.
Critical Event	An unidentified critical event occurs.

Remote Loopback

Figure 31-21 demonstrates the principles of remote loopback. When a local DTE sends non-OAMPDUs to a remote DTE, the remote DTE loops the non-OAMPDUs back to the local DTE, not to the destination addresses of the non-OAMPDUs. This is remote loopback. It helps to obtain the link performance data (including packet loss rate) by comparing the received and transmitted packets.

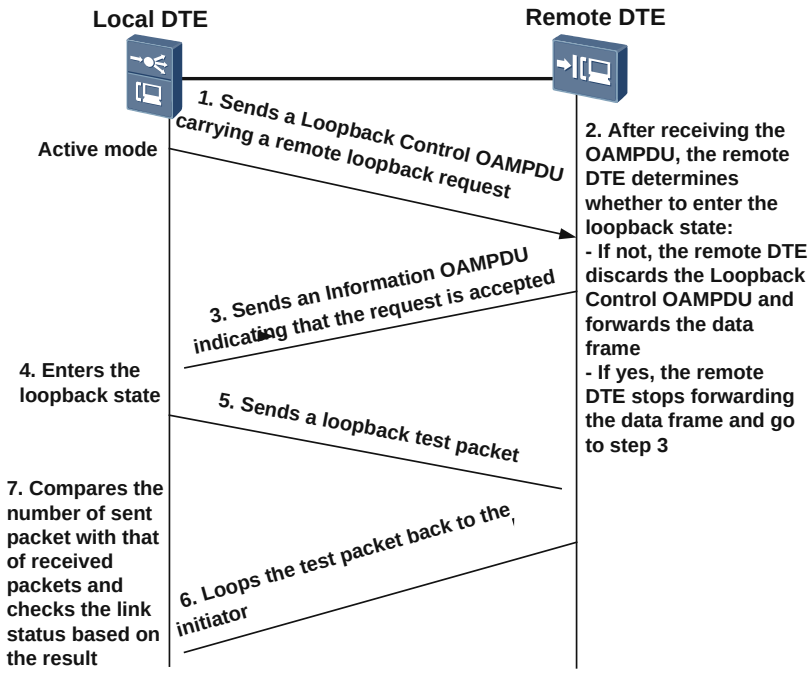
Figure 31-21 Principles of EFM remote loopback



To implement remote loopback, an EFM connection must be established first. After remote loopback is enabled, all data frames except OAMPDUs are dropped and services will be interrupted. This function can be used to detect the link quality before a new network is deployed or after a link fault is rectified to prevent services from being affected.

In remote loopback mode, the local DTE sends testing packets to the remote DTE. The local DTE then computes communication quality parameters, such as the packet loss ratio, of the current link according to the numbers of sent packets and received packets. If the local DTE attempts to stop remote loopback, it sends a message to instruct the remote DTE to disable remote loopback. After receiving the message, the remote DTE disables remote loopback. Figure 31-22 shows the remote loopback process.

Figure 31-22 Remote loopback process



As defined in the IEEE 802.3ah protocol, the EFM loopback initiator needs to block the packets at the Rx end instead of forwarding them to the upper device. This prevents the MAC address forwarding entry of the upper device from being duplicated. On some live networks, service flows are connection-oriented. In this case, packets are not forwarded based on MAC addresses and such a scenario does not involve MAC address duplication. In some other cases, carriers connect the test instrument to the upstream port of an access device, and use the access device as the EFM loopback initiator that forwards the received loopback packets to the upper-layer device. Some access device boards of V800R013C10 or later are developed to meet requirements in such scenarios. These boards support configuration of packet forwarding policy (discarding or forwarding the received loopback packets) in the Rx direction of the EFM loopback initiator.

Link Monitoring

Monitoring Ethernet links is difficult if network performance deteriorates while traffic is being transmitted over physical links. To resolve this problem, the EFM link monitoring function can be used. It can detect data link layer faults in various environments. DTEs that are enabled with link monitoring exchange Event Notification OAMPDUs to monitor links.

After monitoring common link events listed in Table 31-12, the DTE at an end sends OAMPDUs carrying corresponding events to the other end and generates Ethernet OAM link events at local. When the other end receives the OAMPDUs, it generates an event. This facilitates real-time network monitoring and effective network management for maintenance engineers.

Table 31-12 Common link events

Type	Description	Purpose
Errored Symbol	The number of errored symbols	To detect code errors during data

Type	Description	Purpose
Period Event	during a period reaches the threshold.	transmission at the physical layer.
Errored Frame Event	The number of frames during the unit time reaches the preset threshold.	To detect frame errors that occur during data transmission at the MAC sublayer.
Errored Frame Period Event	The number of errored frames in the latest N frames reaches the threshold.	
Errored Frame Seconds Summary Event	The number of errored frames in the latest M seconds reaches the threshold. **NOTE** If an errored frame is generated at a second, this second is called errored frame second.	To detect errored frame seconds that occur during data transmission at the MAC sublayer.

31.5.4 EFM Configuration

Configuring an EFM Remote Loopback

In an Ethernet in the first mile (EFM) remote loopback, all non-OAM protocol data units (PDUs) sent from the local data terminating entity (DTE) to the remote DTE are looped back to the local DTE. The link performance data (such as packet loss rate) can be obtained by comparing the sent and received non-OAM PDUs on the local DTE. This helps to evaluate Ethernet link quality and locate faults.

Context

Before the EFM remote loopback starts, the EFM feature must be enabled on the local DTE and the remote DTE. After the EFM feature is enabled, the EFM mode and the response mode of EFM remote loopback cannot be modified on the port. If these two modes need to be modified, disable the EFM feature.

An EFM remote loopback results in service interruption. Therefore, stop the loopback right after the link performance data is collected.

The following describes how to configure an EFM remote loopback by assuming that the local DTE initiates an EFM remote loopback and the remote DTE responds to the EFM remote loopback.

Procedure

Configure the remote DTE.

1. Run the **efm loopback** *frameid/slotid/portid* **process** command to set the response mode of EFM remote loopback on a port. By default, the response mode of EFM remote loopback on a network-side port is set to **process**, while the response mode of EFM remote loopback on a user-side port is set to **ignore**.

NOTE

Only a few boards can respond to an EFM remote loopback. For these boards, run the above command to set the response mode to **process** or **ignore**, that is, to respond to or ignore the EFM remote loopback initiated by the peer DTE. For boards that cannot respond to an EFM remote loopback, the response mode is defaulted to **ignore** and cannot be modified.

2. Run the **efm oam** *frameid/slotid/portid* **enable** command to enable the EFM feature for the port. By default, the EFM feature is disabled.

Step 1 Configure the local DTE.

1. Run the **efm oam mode** *frameid/slotid/portid* **active** command to set the EFM mode of a port to the default value **active**.
2. Run the **efm oam** *frameid/slotid/portid* **enable** command to enable the EFM feature for the port. By default, the EFM feature is disabled.
3. Run the **efm loopback** *frameid/slotid/portid* **start** command to start the EFM remote loopback. By default, the EFM remote loopback stops.

NOTE

If you want the EFM remote loopback to automatically stop after a period of time, configure the **timeout** parameter to set the loopback timeout period when running the **efm loopback** *frameid/slotid/portid* **start** command.

----End

Result

All packets except OAM PDUs sent from the local DTE to the remote DTE are looped back to the local DTE.

Example

Assume that Ethernet cascade port 0/2/0 on the OLT is connected to upstream port 0/0/1 on the MDU, Ethernet cascade port 0/2/0 on the OLT initiates an EFM remote loopback, and upstream port 0/0/1 on the MDU responds to the EFM remote loopback. To configure an EFM remote loopback on the two ports, do as follows:

1. Configure MDU (MA5616).

```
MDU(config)#efm oam 0/0/1 enable
```

- Configure OLT.

```
OLT(config)#efm oam 0/2/0 enable
OLT(config)#efm loopback 0/2/0 start
```

Follow-up Procedure

After the link performance data is collected, run the **efm loopback** *frameid/slotid/portid* **stop** command on the local DTE to stop the EFM remote loopback.

NOTE

If the timeout period has been set, the EFM remote loopback will automatically stop after the timeout period expires.

Configuring EFM Ethernet Link Monitoring

The Ethernet in the First Mile (EFM) is an operation, administration and maintenance (OAM) feature which provides a real-time link monitoring mechanism. Using this mechanism, events are reported in the case of traffic interruption or link quality degradation so that the network maintenance engineers can monitor Ethernet link quality in real time.

Context

EFM events can be classified into two types: local EFM events and remote EFM events. Local EFM events are detected by the local DTE; remote EFM events are detected by the remote DTE and sent to the local DTE.

- All remote EFM events will be reported as long as the EFM feature is enabled. There is no command that can be used to configure the event reporting function for the remote EFM events.
- For some local EFM events, you need to enable the event reporting function so that the local EFM events can be reported. As listed in the Table 31-13, the local DTE reports the related EFM events and notifies the remote DTE only when the EFM feature and the event reporting function are enabled. If the EFM feature is enabled but the event reporting function is disabled, you can run the **display efm oam event current** *frameid/slotid/portid* **local** command to query the EFM events detected by the local DTE.

Table 31-13 Conditions for detecting and reporting local EFM events

EFM Feature	Event Reporting Function	Detect Local EFM Events or Not	Report Local EFM Events or Not
Enabled	Enabled	Yes	Yes
	Disabled	Yes	No
Disabled	Enabled	No	No
	Disabled	No	No

The local DTE detects only three types of EFM events: link fault event, errored frame event, and errored frame seconds summary event.

- For the errored frame event and errored frame seconds summary event, the event reporting function is disabled by default. You need to run some commands to enable this function. For the specific commands, see the following configuration steps.
- For the link fault event, the event reporting function is enabled by default and cannot be modified.

Procedure

Run the **efm oam** *frameid/slotid/portid* **enable** command to enable the EFM feature for the port. By default, the EFM feature is disabled.

Step 1 (Optional) Configure the event reporting function and related parameters on the local DTE.

When the local DTE need to report the errored frame event or the errored frame seconds summary event and notify the remote DTE, perform this step by running the **efm error-frame** command or **efm error-frame-second** command respectively.

Step 2 Query the configuration results. The following table lists the query commands related to EFM Ethernet link monitoring.

Item	Command
Querying the status of the EFM feature on the local DTE	**display efm oam status** *frameid/slotid/portid* **local** **NOTE** In the command output, the **Admin Status** parameter indicates whether the EFM feature is enabled on the local DTE.
Querying the status of the remote DTE	**display efm oam status** *frameid/slotid/portid* **remote** **NOTE** In the command output, the **Event Support** parameter indicates whether the remote DTE supports sending a notification to the local DTE when the remote DTE detects an EFM event.
Querying the event reporting function and related parameters of different EFM events	**display efm oam event config**
Querying the latest EFM event	**display efm oam event current**
Querying the number of EFM events	**display statistics performance**

----End

Result

In the case of traffic interruption or link quality degradation, the access device (local DTE) reports events based on the configurations.

Example

Assume that Ethernet cascade port 0/2/0 on the OLT is connected to upstream port 0/0/1 on the MDU, both OLT and MDU can receive the EFM events sent by each other and can detect EFM events, and the detection interval and other thresholds are defaulted. To monitor the quality of this Ethernet link section, do as follows:

- OLT configurations:

```
OLT(config)#efm oam 0/2/0 enable
OLT(config)#efm error-frame 0/2/0 notification enable
OLT(config)#efm error-frame-second 0/2/0 notification enable
```

- MDU configurations:

```
MDU(config)#efm oam 0/0/1 enable
MDU(config)#efm error-frame 0/0/1 notification enable
MDU(config)#efm error-frame-second 0/0/1 notification enable
```

```
MDU(config)#efm error-frame-period 0/0/1 notification enable
```

NOTE

Of the OLT and MDU, only the MDU can detect the errored frame period event, and only MDUs of some models can detect this event. For details, see the *Feature Specifications* in related MDU product documents.

31.5.5 EFM Maintenance and Diagnosis

After EFM is enabled, the access device uses 0x2f00000a Ethernet OAM link events to implement remote failure indication and link monitoring. According to event parameters, the event type (such as link fault and errored frame) and location (local or remote) can be determined.

- The access device can receive all types of EFM events sent by the remote device and then report 0x2f00000a Ethernet OAM link events at local.
- The access device can only detect some types of EFM events, report 0x2f00000a Ethernet OAM link events at local, and uses the OAMPDU to notify the remote end. Types of EFM events can be detected at local are related to the hardware. For the hardware supported, see "Specifications."

After the EFM connection is established, the remote DTE status can be queried by the **display efm oam status** command.

The statistics of EFM OAM packets that are received and sent by local TDE can be queried by the **display efm oam statistics** command.

31.6 PM (Y.1731)

Performance monitoring (PM) is defined in ITU-T Y.1731. PM, as an OAM feature, is used to test the performance of Ethernet links and evaluate the network quality.

31.6.1 PM Introduction

This topic describes the benefits and functions of performance monitoring (PM) defined in ITU-T Y.1731.

Benefits

As the multi-play service (for example, IPTV) prevails, more and more difficulties in Ethernet OAM emerge. The following uses the triple-play service as an example to explain the difficulties facing carriers.

- IPTV: Video streams are sensitive to packet loss and delay. The loss of a packet affects dozens of pictures, causing frame freezing and impairing user experience.
- VoIP: Audio streams are sensitive to packet loss and delay. Packet loss or overlong delay results in intermittent voices; delay variations result in vague voices, which is intolerable to users.
- High-speed Internet (HSI): Packet loss or delay occurring in the data streams may cause user authentication failures or user connection failures, which increases user complaints.

As a result, minimizing user complaints (caused by network quality degradation) and improving fault locating efficiency have become carriers' top concerns.

Besides, in the case of national broadband, carriers charge retail service providers (RSPs) based on the level of network quality when selling bandwidth to the RSPs. Some carriers even charge end users based on the level of network quality. Therefore, carriers need to monitor the network quality and provide the service level agreement (SLA) as the accounting basis.

ITU-T Y.1731 defines a series of methods for monitoring Ethernet link performance. These methods help carriers collect link performance data in real time and fast diagnose the network performance faults. ITU-T Y.1731 also provides carriers with technical means for outputting the SLA reports.

Functions

The following table lists the three PM functions supported by access devices.

Function	Description
Single-ended ETH-LM	Single-ended Ethernet loss measurement (ETH-LM) is used to collect the packet loss rate of an Ethernet link based on a count of transmitted and received data frames between a pair of MEPs.
Two-way ETH-DM	Two-way Ethernet delay measurement (ETH-DM) is used to collect the delay, delay variation of an Ethernet link.
Single-ended ETH-SLM	Single-ended Ethernet synthetic loss measurement (ETH-SLM) is a mechanism to measure frame loss using synthetic frames, rather than data traffic. A number of synthetic frames are sent and received, and the number of those that are lost is hence calculated. This can be treated as a statistical sample, and used to approximate the frame loss ratio of data traffic.

Both single-ended ETH-LM, two-way ETH-DM and single-ended ETH-SLM support the following modes:

- On-demand mode: In this mode, performance data is collected in real time. On-demand single-ended ETH-LM and two-way ETH-DM are used for diagnosing network quality faults. They can be started through the CLI or NMS, and related measurement results are displayed on the CLI or NMS.
- Periodic mode: In this mode, performance data is collected and reported periodically. Periodic single-ended ETH-LM and two-way ETH-DM are used by carriers to output periodic SLA reports. They can be started only through the NMS.

31.6.2 PM Networking Application

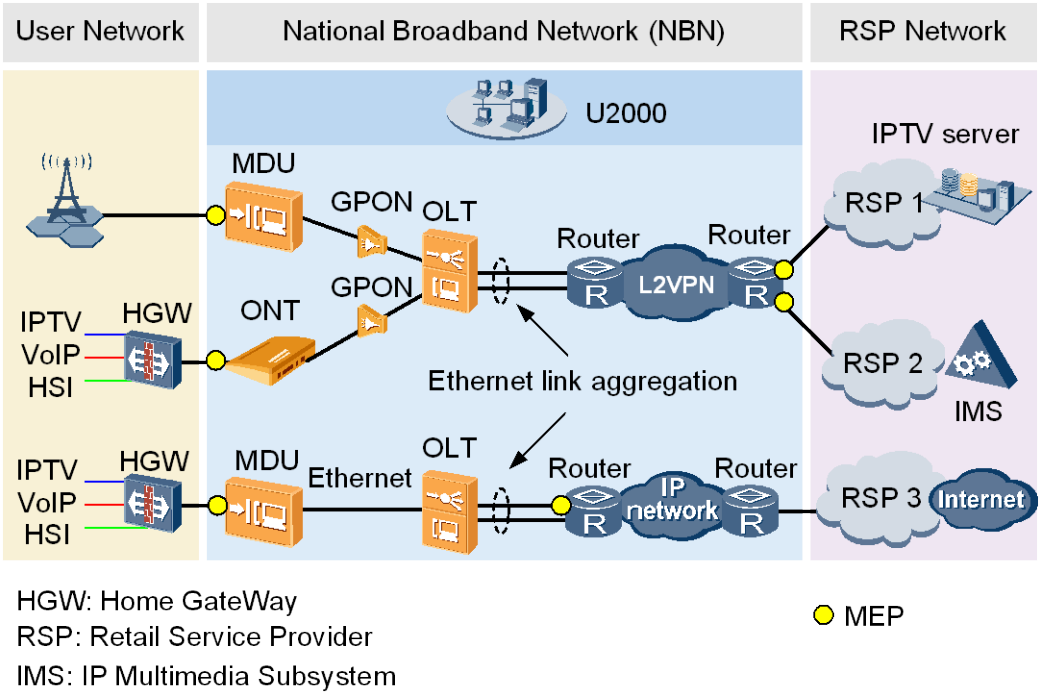
Performance monitoring (PM) is mainly used on the national broadband (NBB) network to monitor and evaluate carriers' service quality. In addition, some carriers deploy maintenance association end points (MEPs) on their networks. These MEPs help carriers test link performance section by section along a network link, achieving fast fault demarcation and fault locating.

The following describes the typical MEP deployment on common networks.

NBB Network

Figure 31-23 shows the typical MEP deployment on the NBB network to implement PM according to ITU-T Y.1731. The user ports on the MDUs and ONTs mark the boundary between the NBB network and the user network; the network-side ports on routers mark the boundary between the NBB network and the retail service provider (RSP) network. MEPs deployed on the boundaries of the NBB network collect the performance data about the Ethernet links between 2 MEPs. The collected performance data is used for network quality evaluation. ITU-T Y.1731 is a Layer 2 network protocol, and the Ethernet protocol packets defined by ITU-T Y.1731 cannot traverse an IP network. If MEPs need to be deployed on the network-side ports on the routers at the boundary of the NBB network, a Layer 2 VPN tunnel needs to be established between the routers.

Figure 31-23 Typical MEP deployment on the NBB network



NOTE
Currently, no ONT supports Ethernet synthetic loss measurement (ETH-SLM) defined in ITU-T Y.1731.

FTTB/FTTC Network

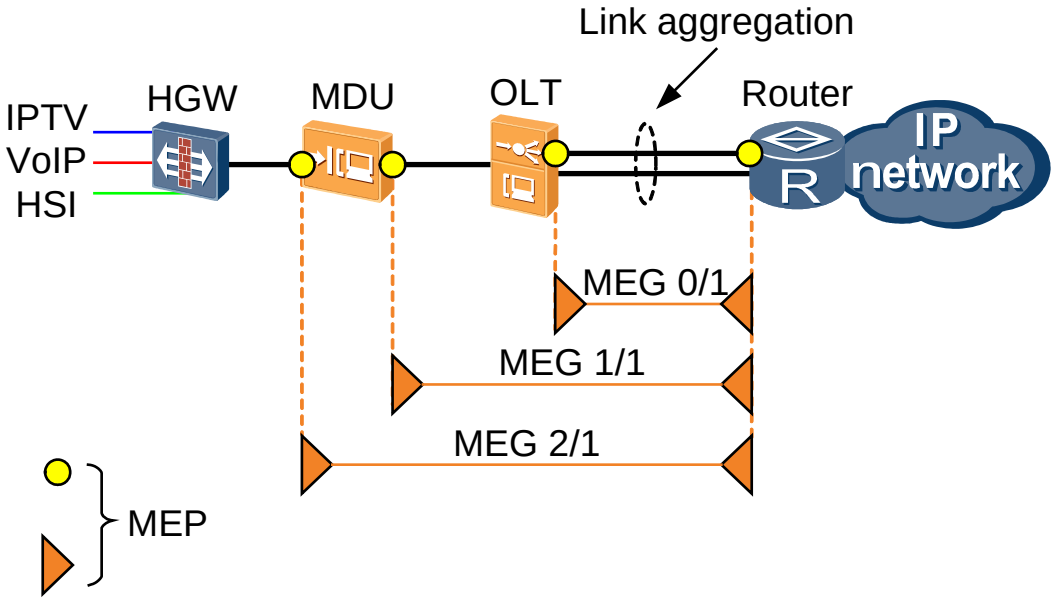
Figure 31-24 shows the MEPs deployed on the user port and upstream port on the MDU, upstream link aggregation group of the OLT, and downstream link aggregation group of the router on a fiber to the building (FTTB) or a fiber to the curb (FTTC) network. These MEPs help test the link performance in each maintenance entity group (MEG).

For example, if a user reports that the voice is intermittent during the call, maintenance engineers perform single-ended Ethernet loss measurement (ETH-LM) section by section along the network link to diagnose the specific network section in which the user's voice service stream encounters packet loss. As shown in Figure 31-24, maintenance engineers check whether packets are lost in MEG 2/1. If no packet is lost, the link between the user port

on the MDU and the router runs normally. If packets are lost, maintenance engineers continue to check whether packets are lost in MEG 1/1.

- If packets are lost in MEG 1/1, maintenance engineers check whether packets are lost in MEG 0/1.
- If no packet is lost in MEG 1/1 but packets are lost in MEG 2/1, it indicates that packets are lost on the MDU.

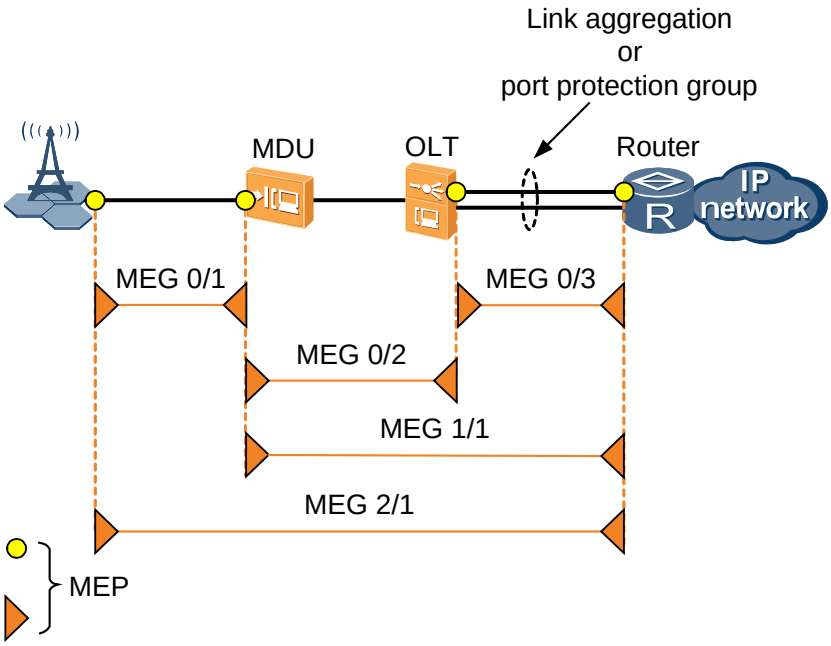
Figure 31-24 Typical MEP deployment on an FTTB/FTTC network



FTTM Network

Fiber to the mobile base station (FTTM) networks have a high requirement on link reliability. Therefore, port protection groups or link aggregation groups (active/standby mode) are created on the links between the OLT and router on an FTTM network. Figure 31-25 shows the MEPs deployed on the upstream port on the mobile base station, user port on the MDU, upstream port protection group or link aggregation group of the OLT, and user-side port protection group or downstream link aggregation group of the router. These MEPs help test the link performance in each MEG.

Figure 31-25 Typical MEP deployment on an FTTM network



FTTH Network

The MEP is deployed on the LAN port or the WAN port on the ONT, upstream link aggregation group of the OLT, and downstream link aggregation group of the router. These MEPs help test the link performance in each MEG.

- Figure 31-26 shows the up MEP deployed on the LAN port of a bridging ONT.
- Figure 31-27 shows the down MEP deployed on the WAN port of a routing ONT.

Figure 31-26 Typical MEP deployment on an FTTH network with bridging ONT

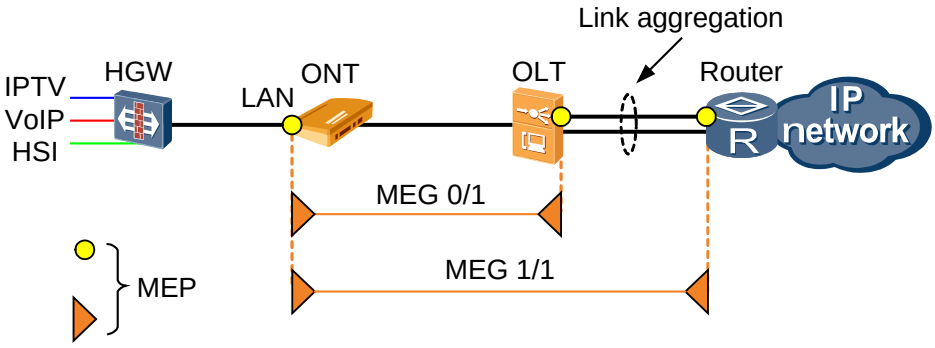


Figure 31-27 Typical MEP deployment on an FTTH network with routing ONT

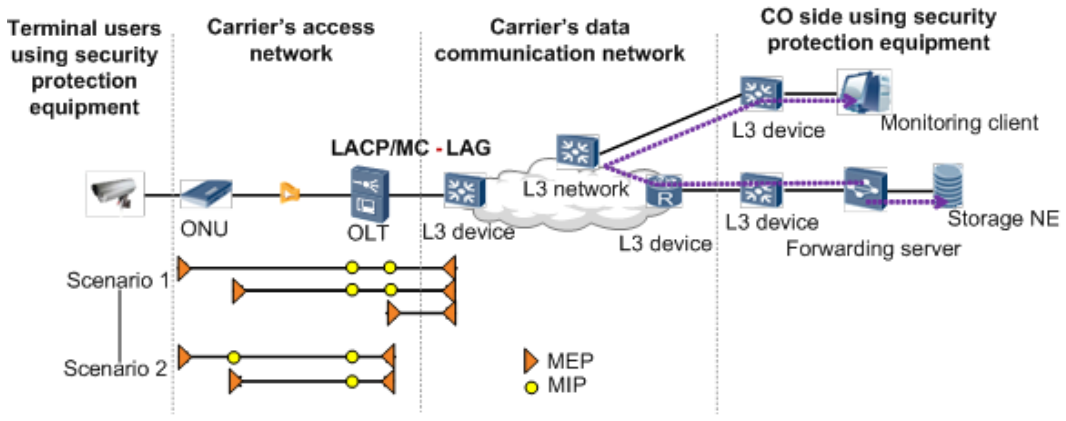
In FTTH scenario, the performance monitoring function of ONT is configured through the CLI of OLT. Then OLT issues the configuration to ONT through the optical network terminal management and control interface (OMCI) channel. ONT can only function as the responding end.

FTTO Video Monitoring Network

On-demand video monitoring is preferred for fault demarcation, as shown in Figure 31-28.

- Scenario 1. MEGs are deployed on the Layer 3 UNI and ONU UNI to check continuity and check packet loss, delay, and jitter in one direction. In this scenario, the protection network between the upstream port on the OLT and Layer 3 device needs to be considered, including single-homing Link Aggregation Control Protocol (LACP) protection and dual-homing multi-chassis link aggregation group (MC-LAG) protection.
- Scenario 2. MEGs are deployed on the OLT NNI and ONU UNI to check continuity and check packet loss, delay, and jitter in one direction.

Figure 31-28 Typical MEP deployment on an FTTO video monitoring network



31.6.3 PM Basic Concepts

This topic introduces the basic concepts of performance monitoring (PM) defined in ITU-T Y.1731.

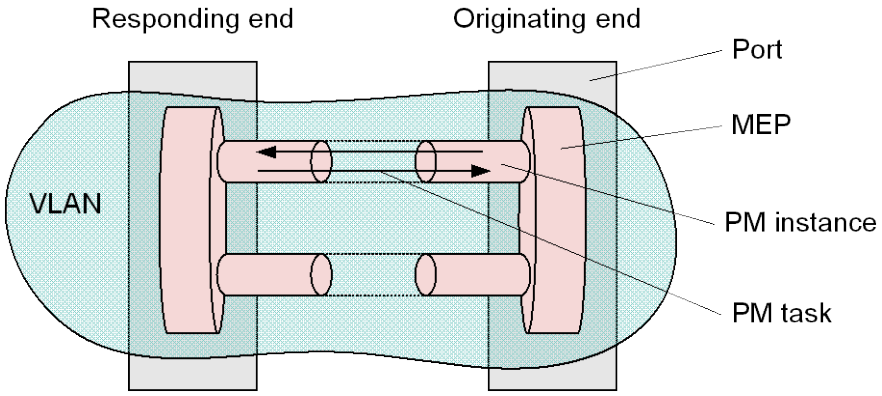
Some concepts are detailed in ITU-T Y.1731 and IEEE 802.1ag. When access devices implement PM, some concepts, such as the maintenance association end point (MEP) and maintenance association (MA) are used. For details, see 31.4.3 CFM Basic Concepts and 31.3 Differences in Implementing Y.1731 and 802.1ag on Access Device.

Besides, Huawei introduces some concepts (such as the PM instance and PM task) that are not defined in the protocols. The following table lists the items that need to be configured during PM deployment. The relationships between the configured items are shown in Figure 31-29.

Item	Function
MEP	Used to demarcate the Ethernet link to be monitored. MEPs can be divided into originating and responding ends based on the type of Ethernet protocol packets sent by them. • Originating end: sends the Ethernet protocol request packets and receives the Ethernet protocol response packets sent by the responding end. By comparing the request and response packets, the system calculates the packet loss rate, number of lost packets, and the delay occurring on the link between the 2 MEPs. • Responding end: replies to the originating end with Ethernet protocol response packets. The responding end does not calculate or save the performance test results.
PM instance	Used to select the PM object on the Ethernet link between 2 MEPs. PM instances are identified by the test ID. The following parameters are provided to differentiate between different PM instances. • **mep**, **remote-mep**, and **remote-mep-mac**: specify the two ends of the Ethernet link to be monitored. Two MEPs exchange Ethernet protocol packets with each other to collect performance data. • **receive-8021p** and **send-8021p**: specify the priority of Ethernet protocol packets sent and received by the MEP. The priority of the Ethernet protocol packets must be the same as that of the to-be-monitored service flow. In Multi-priority Convergence scenario, to calculate all the packets received by or sent from the MEP according to Huawei's proprietary protocol, set these parameters to **unaware**. Parameter value **unaware** is valid to Ethernet loss measurement (ETH-LM), Ethernet delay measurement (ETH-DM), and Ethernet synthetic loss measurement (ETH-SLM). • **backward-mac**, **onward-mac**, **update-backward-mac**, and **update-onward-mac**: defined by Huawei's proprietary protocol. These 4 parameters are used to differentiate between service flows when an ETH-LM or ETH-DM test is performed in the point-to-multipoint connection (N:1 VLAN) scenario. These 4 parameters do not need to be configured when an Ethernet synthetic loss measurement ETH-SLM is performed or when an ETH-LM or ETH-DM test is performed in the point-to-point connection (1:1 VLAN) scenario.
PM task	Used to select one of the PM instances configured in the system for an LM or delay measurement (DM) test. PM instances are configured manually and do not trigger the Ethernet protocol packet exchange between the MEPs at both ends of the link. The 2 MEPs normally exchange Ethernet protocol packets only when the originating and responding ends use the same PM instance for performing the same PM

Item	Function
	task. • **pm lm send** and **pm dm send**: used to configure the interval for sending Ethernet protocol packets (used for LM and DM tests) and configure the number of Ethernet protocol packet sending attempts on the originating end. After these 2 parameters are configured, the originating end starts to send Ethernet protocol request packets. • **pm lm receive**, **pm dm receive**, and **pm slm receive**: used to configure whether the responding end responds to the requests initiated by the originating end.

Figure 31-29 Items that need to be configured during PM deployment



31.6.4 PM Principles

Performance monitoring (PM) defined in ITU-T Y.1731 cannot be used in some application scenarios. Therefore, Huawei broadens the PM application scenarios based on Huawei's proprietary protocol. The following introduces how to implement Ethernet link PM on an FTTB/FTTC network with typical VLAN and QoS planning.

Access devices implement single-ended Ethernet loss measurement (ETH-LM), two-way Ethernet delay measurement (ETH-DM) and single-ended Ethernet synthetic loss measurement (ETH-SLM) according to ITU-T Y.1731. For basic implementation principles, see ITU-T Y.1731.

Performance Monitoring with Typical VLAN Planning

Single-ended Ethernet loss measurement (ETH-LM) and two-way Ethernet delay measurement (ETH-DM) defined in ITU-T Y.1731 are based on the point-to-point (P2P) connection. That is, a maintenance association end point (MEP) receives packets only from a remote MEP. However, in actual networking applications, the VLAN planning is very flexible, which means that a MEP receives packets from multiple remote MEPs. This networking mode is a type of point-to-multipoint (P2MP) connection mode. The following introduces how to implement single-ended Ethernet loss measurement (ETH-LM) and two-way Ethernet delay measurement (ETH-DM) in P2P and P2MP connection scenarios with typical VLAN planning.

P2P Connection (1:1 VLAN Scenario)

Figure 31-30 P2P network for enterprise users

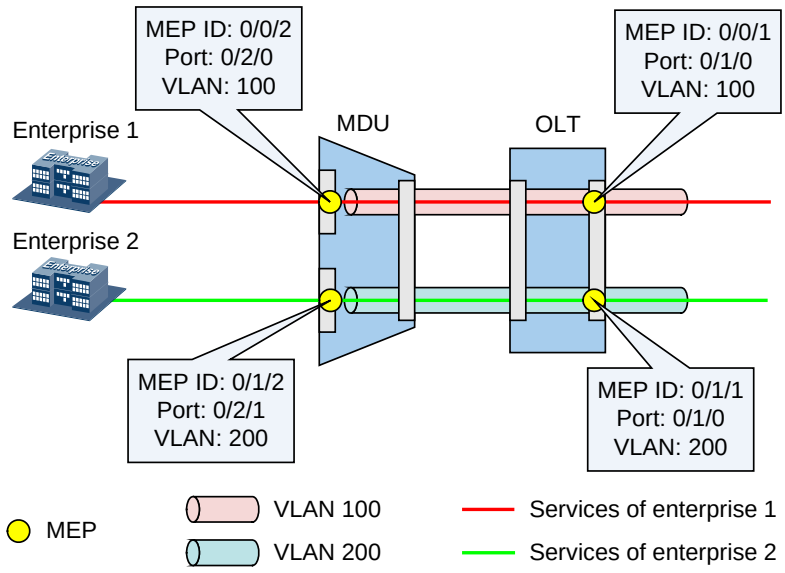


Figure 31-30 shows the typical 1:1 VLAN planning. The MDU translates the VLAN of the enterprise user's service flow to the service VLAN (S-VLAN). Then the OLT transparently transmits the VLAN. The following conditions must be met during the S-VLAN planning.

- Different enterprise users connected to the same OLT must be configured with different S-VLANs.
- Packets sent by any 2 MDUs connected to the same OLT cannot carry the same S-VLAN tag.

Assume that enterprise users' services in Figure 31-30 are transmitted upstream by the SPUF board on the OLT. If maintenance engineers need to check the performance of the link between the MDU user port and the OLT upstream port, deploy the MEPs on the MDU user port and the OLT upstream port as shown in Figure 31-30. All the packets received on MEP 0/0/1 are sent from MEP 0/0/2. Such a connection is a P2P connection. For a P2P connection, when a PM instance is added, the MEPs at both ends of the link and the priorities of to-be-test services must be specified.

P2MP Connection (N:1 VLAN Scenario)

Figure 31-31 P2MP network for home users

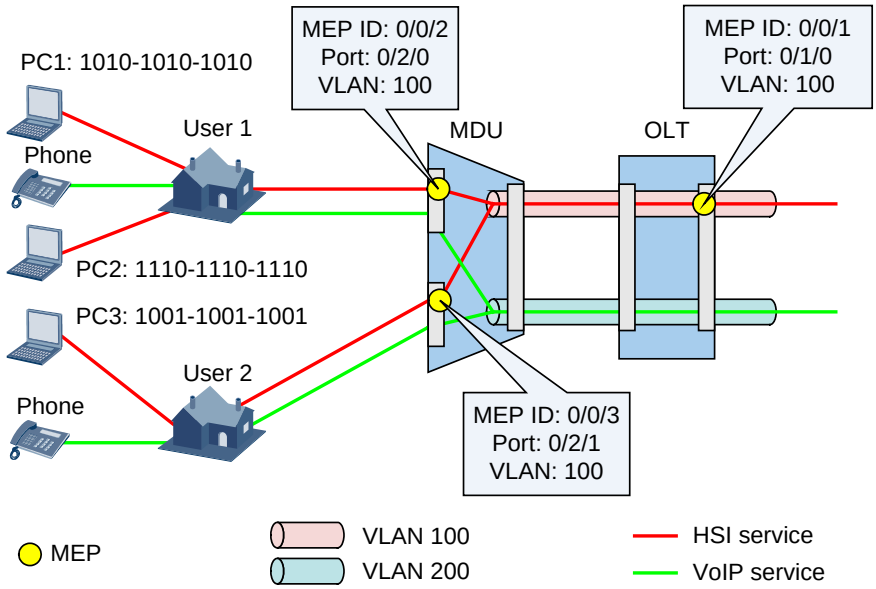


Figure 31-31 shows the typical N:1 VLAN planning. The MDU translates the VLAN of the home user's service flow to the S-VLAN. Then the OLT transparently transmits the VLAN. The following conditions must be met during the VLAN planning.

- Customer VLANs (C-VLANs) are used to differentiate between home users' service flows.
- The MDU translates C-VLANs to S-VLANs. Packets of the same type are transmitted upstream through the same S-VLAN.

Assume that home users' services in Figure 31-31 are transmitted upstream through the SPUF board on the OLT. If maintenance engineers need to measure the packet loss rate of user 1's HSI services on an FTTx network, deploy the MEPs on the MDU user port and the OLT upstream port as shown in Figure 31-31. All data services on the MDU are transmitted upstream through the same S-VLAN. All the packets received on MEP 0/0/1 are sent from MEP 0/0/2 and MEP 0/0/3. Such a connection is a P2MP connection. In this case, MEP 0/0/1 cannot differentiate between packets from MEP 0/0/2 and those from MEP 0/0/3. Therefore, the packet loss rate cannot be accurately measured on the link between MEP 0/0/1 and MEP 0/0/2. To solve this issue, Huawei defines 2 concepts, onward MAC and backward MAC, in its proprietary protocol. Onward MAC and backward MAC are used to identify P2P service flows on a P2MP connection network.

Essentially, onward MAC and backward MAC are MAC addresses of user terminals. Each MAC address is globally unique. Therefore, the MEP where service flows converge can identify the source of the packets based on the MAC address.

- Backward MAC: used to configure the PM instance at the access end (for example, the MDU user port in Figure 31-31) of the service flows.
- Onward MAC: used to configure the PM instance at the convergence end (for example, the OLT upstream port in Figure 31-31) of the service flows.

Onward MAC and backward MAC can be manually specified or randomly selected by the device.

- When maintenance engineers need to measure the packet loss rate of a specific user's service flows on a link, use the **onward-mac** and **backward-mac** parameters to manually specify the MAC address of the user terminal.
- When maintenance engineers need to measure the packet loss rate of the whole links, use the **update-onward-mac** and **update-backward-mac** parameters to make the device randomly select the service flows from a user terminal as the test sample.

Assume that, in Figure 31-31, the priority of the upstream and downstream packets of user 1's and user 2's HSI services is 0; the MEP on the OLT functions as an originating end; the MEP on the MDU functions as a responding end. The following table lists PM instance configurations (using Figure 31-31 as an example).

Test Object	PM Instance Configuration on the MDU	PM Instance Configuration on the OLT
Data services of user 1	**pm instance mep 0/0/2 remote-mep 1 receive-8021p 0 send-8021p 0 backward-mac 1010-1010-1010**	**pm instance mep 0/0/1 remote-mep 2 receive-8021p 0 send-8021p 0 onward-mac 1010-1010-1010**
HSI services of user 1	**pm instance mep 0/0/2 remote-mep 1 receive-8021p 0 send-8021p 0 update-backward-mac**	**pm instance mep 0/0/1 remote-mep 2 receive-8021p 0 send-8021p 0 update-onward-mac**
HSI services of user 2	**pm instance mep 0/0/3 remote-mep 1 receive-8021p 0 send-8021p 0 update-backward-mac**	**pm instance mep 0/0/1 remote-mep 3 receive-8021p 0 send-8021p 0 update-onward-mac**

On a P2MP connection network, the access device differentiates between service flows according to the MAC addresses of user terminals. The following requirements must be met during PM instance deployment.

- The MAC address learning function is not disabled at the access end (for example, the MDU in Figure 31-31).
- The maximum number of learnable dynamic MAC addresses on the service port must be greater than or equal to the actual number of MAC addresses on the service port to ensure that all the source MAC addresses of packets sent by user terminals can be learnt by the access device. Otherwise, the access device may regard the user packets as unknown unicast packets and discard them. Therefore, the test results may be inaccurate.
- On an FTTB/FTTC network, if an up MEP is configured on the MDU service board, 1:1 VMAC can be enabled on the MDU, while N:1 VMAC address cannot.
- The access device implements single-ended Ethernet loss measurement (ETH-LM) on a P2MP connection network according to Huawei's proprietary protocol, and therefore the access devices interconnect only with Huawei devices. When the peer MEP is on a device which is not manufactured by Huawei, ETH-SLM can be used to collect the packet loss rate on a P2MP connection network.

- If the performance of the service flow needs to be monitored, the traffic classification parameter of the **service-port** must be created based on **user-vlan** instead of **user-8021p**, **user-8021p-list** or **user-encap**.

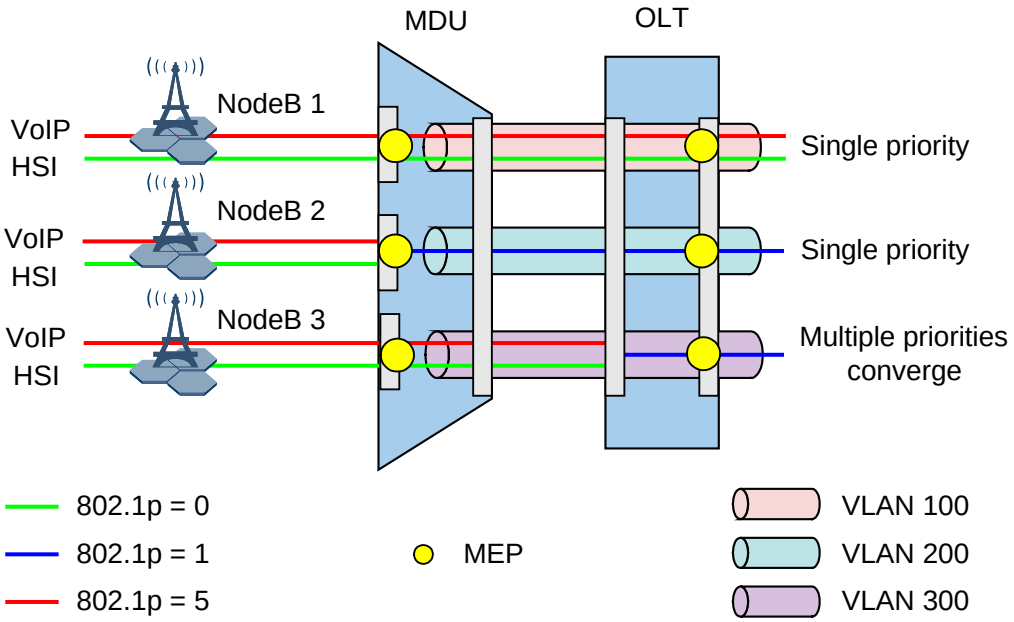
Performance Monitoring with Typical QoS Planning

As defined in ITU-T Y.1731, the priority of Ethernet protocol packets used for single-ended Ethernet loss measurement (ETH-LM), two-way Ethernet delay measurement (ETH-DM) and single-ended Ethernet synthetic loss measurement (ETH-SLM) must be configurable to make the priority of Ethernet protocol packets equal to that of to-be-tested packets. If the priority of to-be-tested packets changes during packet forwarding, the priority of Ethernet protocol packets must be changed accordingly. The following describes the PM configuration and implementation principles on an FTTM point-to-point (P2P) network with different QoS plans.

Single Priority and Multi-priority Convergence

A service flow may contain packets with different priorities. Generally, the system chooses packets with the same priority for PM tests. This belongs to the single priority scenario (for example, NodeB 1 and NodeB 2 in Figure 31-32). However, in actual networking applications, carriers may modify packet priorities based on the QoS planning. Therefore, a service flow on a MEP may contain packets with different priorities, but this service flow on another MEP contains packets with the same priority. This belongs to the multi-priority convergence scenario (for example, NodeB 3 in the Figure 31-32).

Figure 31-32 Single priority and multi-priority convergence



Assume that the 802.1p field is used to differentiate between service flows. Carriers configure different priority processing policies for different NodeBs, as shown in Figure 31-32.

- NodeB 1: The MDU and OLT copy the 802.1p priority of user-side packets as the priority of upstream packets and copy the 802.1p priority of network-side packets as the priority of downstream packets.

- NodeB 2: The MDU specifies the 802.1p priority of upstream packets to 1, and the OLT copies the 802.1p priority of user-side packets as the priority of upstream packets. The MDU and OLT copy the 802.1p priority of network-side packets as the priority of downstream packets.
- NodeB 3: The MDU copies the 802.1p priority of user-side packets as the priority of upstream packets, and the OLT specifies the 802.1p priority of upstream packets to 1. The MDU and OLT copy the 802.1p priority of network-side packets as the priority of downstream packets.

As shown in Figure 31-32, MEPs are deployed on the MDU user port and the OLT upstream port to monitor the performance of Ethernet links.

- For NodeB 1 and NodeB 2, MEPs at both ends of a link perform PM tests based on packets with the same priority in upstream and downstream directions. This belongs to the single priority scenario.
- For NodeB 3, a service flow on the MEP of the MDU contains packets with different priorities, while the service flow on the MEP of the OLT contains packets with the same priority. This belongs to the multi-priority convergence scenario. However, this scenario is not defined in ITU-T Y.1731. Therefore, the access device implements ETH-LM, ETH-DM and ETH-SLM of the multi-priority convergence scenario based on the Huawei's proprietary protocol. As defined in Huawei's proprietary protocol, the priority of Ethernet protocol packets on the MEP containing packets with different priorities is set to **unaware** so that all the packets sent to or received by the MEP are counted during statistics measurement. For example, to perform LM between the user port of the MDU and the upstream port of the OLT, set the priority of Ethernet protocol packets of the PM instance corresponding to NodeB 3 to **unaware** on the MDU.

Assume that the 802.1p priority of downstream packets is set to 0; the MEP on the OLT functions as an originating end; the MEP on the MDU functions as a responding end. The following table lists the priority configurations of Ethernet packets in PM instances. The priorities of services on NodeB 2 and NodeB 3 are modified to the same priority. Therefore, the system can monitor only the performance of all the services, but cannot monitor the performance of a specific service.

Test Object	Priority Configurations on the MDU	Priority Configurations on the OLT
VoIP services of NodeB 1	**receive-8021p 0** **send-8021p 5**	**receive-8021p 5** **send-8021p 0**
HSI services of NodeB 1	**receive-8021p 0** **send-8021p 0**	**receive-8021p 0** **send-8021p 0**
Services of NodeB 2	**receive-8021p 0** **send-8021p 1**	**receive-8021p 1** **send-8021p 0**
Services of NodeB 3	**receive-8021p 0** **send-8021p unaware**	**receive-8021p 1** **send-8021p 0**

In the multi-priority convergence scenario, the priority after multiple priorities converge is used as the priority of Ethernet protocol packets, which is different from the priority of the service flow. Therefore, PM results may be incorrect.

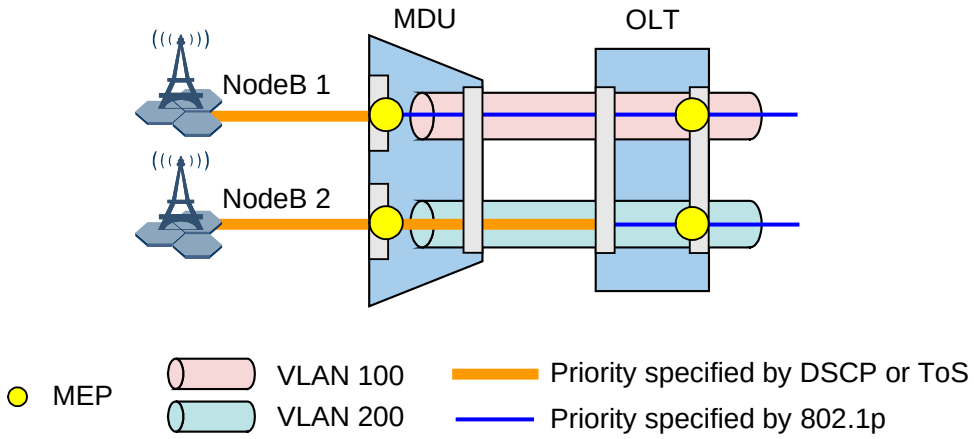
For example, if PM is performed on NodeB 3 in Figure 31-32, service packets forwarded by the MDU to the upstream direction are with different priorities, and the 802.1p priority of

Ethernet protocol packets sent by the MEP on the MDU is 1. The priority of the Ethernet protocol packets is not the same as that of any service packets. In this case, the sequence of Ethernet protocol packets and service packets received on the OLT may change, which may be different from that of Ethernet protocol packets and service packets sent by the MDU. Consequently, the PM results may be inaccurate.

In the multi-priority convergence scenario, to ensure accurate PM results, ensure that the originating end, responding end, and devices in between map packets with different priorities into the same queue for scheduling so that the sequence of Ethernet protocol packets and service packets is not changed during scheduling.

Setting the 802.1p Priority Based on the ToS and DSCP Priorities

Figure 31-33 Setting the 802.1p priority based on the ToS and DSCP priorities



Assume that the ToS or DSCP field is used to differentiate between service flows on NodeBs, carriers configure different priority processing policies for different NodeBs, as shown in Figure 31-33.

- NodeB 1: The MDU sets the 802.1p priority of upstream packets based on the ToS or DSCP priority of the IP packets, and the OLT copies the 802.1p priority of user-side packets as the priority of upstream packets. The MDU and OLT copy the 802.1p priority of network-side packets as the priority of downstream packets.
- NodeB 2: The MDU forwards upstream packets based on the priority of the IP packets, and the OLT sets the 802.1p priority of the upstream packets based on the ToS or DSCP priority of the IP packets. The MDU and OLT copy the 802.1p priority of network-side packets as the priority of downstream packets.

As shown in Figure 31-33, MEPs are deployed on the MDU user port and the OLT upstream port to monitor the performance of Ethernet links.

- For NodeB 1, MEPs at both ends of a link perform PM tests based on packets with the same 802.1p priority in upstream and downstream directions (the 802.1p priority is the one that the MDU sets based on the ToS or DSCP priority). This belongs to the single-priority scenario.
- For NodeB 2, the MEP on the MDU performs PM tests based on the packets with the ToS or DSCP priority in upstream and downstream directions. The priority of the Ethernet protocol packets sent by the MDU is the same as the ToS or DSCP priority of service packets. The priority of service packets received on the OLT is the 802.1p

priority that the OLT sets based on the ToS and DSCP priority. However, Ethernet protocol packets are non-IP packets and they do not carry the ToS or DSCP field. When the OLT modifies the packet priority, the OLT cannot set the priority of Ethernet protocol packets to be the same as that of service packets. In this case, packet disorder occurs among Ethernet protocol packets and service packets, and consequently the PM results are incorrect. Therefore, PM is not supported in this case.

Assume that the 802.1p priority of upstream service packets is set to 1 and that of downstream service packets is set to 0 based on the ToS or DSCP priority of IP packets; the MEP on the OLT functions as an originating end; the MEP on the MDU functions as a responding end. The following table lists the priority configurations of Ethernet packets in PM instances.

Test Object	Priority Configurations on the MDU	Priority Configurations on the OLT
Services of NodeB 1	receive-8021p 0 send-8021p 1	receive-8021p 1 send-8021p 0
Services of NodeB 2	Not supported	Not supported

31.6.5 PM Configuration

This topic describes how to configure ITU-T Y.1731-defined performance monitoring (PM) using the CLI. The CLI supports only the configuration of on-demand PM tests, such as Ethernet loss measurement (ETH-LM), Ethernet delay measurement (ETH-DM) and Ethernet synthetic loss measurement (ETH-SLM). Periodic PM tests are configured only on the NMS. For the detailed operations, see the U2000 configuration guide.

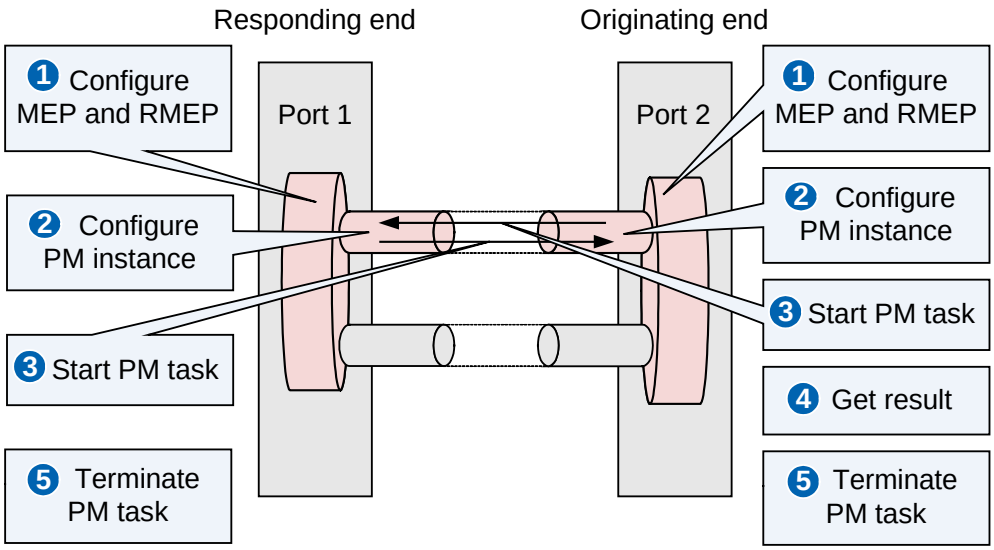
Prerequisites

Service configurations (such as VLAN and service flow configurations) are completed, and user-side services run normally.

Configuration Procedure

Figure 31-34 shows the overall procedure of configuring and performing a PM task on the originating and responding ends.

Figure 31-34 PM configuration procedure



As shown in Figure 31-34, the configuration items and steps on the originating end are similar to those on the responding end. The following table lists the related configuration commands and precautions.

Step	Command	Precautions
1. Configure maintenance association end points (MEPs).	**cfm md**	Parameter **name-format** must be set to **no-name**.
	cfm ma	The maintenance association (MA) ID must be the same as the maintenance entity group (MEG) ID, and parameter **name-format** must be set to **icc-based**.
	cfm ma vlan	If the MEP is configured on the SPUF board, the VLAN associated with the MA must be the same as that configured in **vlantag1** of the MEP in the MA. **NOTE** For a MEP on the ONT, do not configure this command.
	cfm ma meplist	The MEP list must contain the MEP and RMEP. **NOTE** For a MEP on the ONT, do not configure this command.
	cfm mep	For detailed limitations on MEP configurations, see Feature Dependency and Limitation.
2. Configure PM instances.	**pm instance**	• The parameters that need to be configured varies with the VLAN and QoS planning. For details, see Performance Monitoring with Typical VLAN Planning and Performance Monitoring with Typical QoS Planning. • You are advised to set the test ID on the originating end to be the same as that on the responding end to facilitate future management

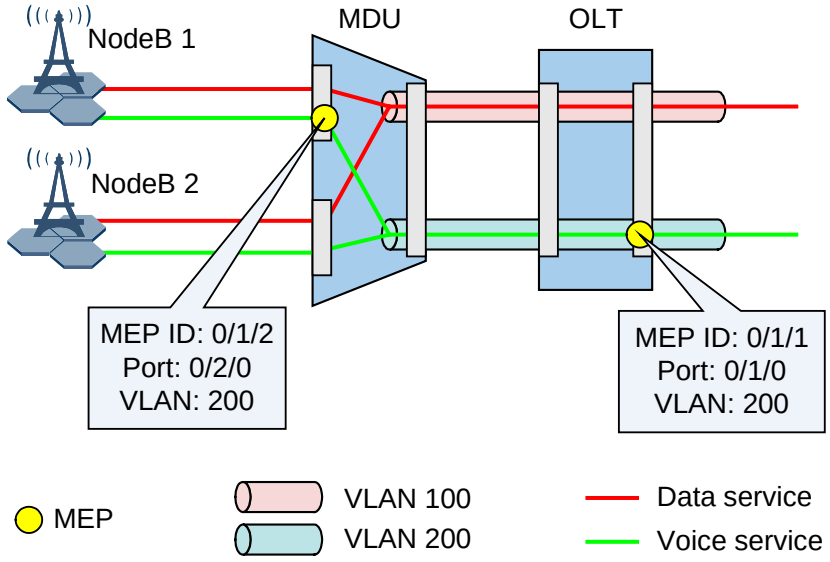
Step	Command	Precautions
		and maintenance if these 2 test IDs belong to the same PM instance. • The **display pm instance** command can be used to query the configured PM instances in the system.
3. Start PM tasks.	ETH-LM 1. Responding end: **pm lm receive** 2. Originating end: **pm lm send** ETH-DM 1. Responding end: **pm dm receive** 2. Originating end: **pm dm send** ETH-SLM 1. Responding end: **pm slm receive**	• The number of started PM tasks must comply with that of PM tasks that can be supported by the devices where the originating and responding ends are configured. • For an LM test, if one of the 2 MEPs (belonging to the same PM instance) is set as the originating end on a device, the other MEP cannot be set as the responding end on the same device. Similarly, if one of the 2 MEPs is set as the responding end on a device, the other MEP cannot be set as the originating end on the same device. • After a PM task is started on the responding end, the responding end sends the response packets to the peer MEP only when the test ID configured for the PM task on the responding end matches the MEP configured on the originating end. • The **display pm task** command can be used to query the started PM tasks on the device. If a PM task is interrupted abnormally, but the timeout interval for the PM task does not expire, this PM task can be queried by running the **display pm task** command. • The **display pm oam pdu statistics** command can be used to query the Ethernet protocol packets that are transmitted and received by the device.
4. Obtain the test results.	**display pm statistics**	You can use any of the following methods to obtain the test results on the originating end. • The test results will be displayed on the CLI in real time after a PM task is started. • You can run the **display pm statistics** command to query the test results after the test is completed.
5. Stop a PM task.	ETH-LM 1. Originating end: - 2. Responding end: **undo pm lm receive** ETH-DM 1. Originating end: - 2. Responding end: **undo pm dm receive**	A PM task on the originating end will be stopped if any of the following conditions is met. • The maintenance engineer presses **Ctrl+C** to manually stop the PM task. • The PM task is completed successfully, or the timeout interval for a PM task has expired so that the system automatically stops the PM task.

Step	Command	Precautions
	ETH-SLM 1. Responding end: **undo pm slm receive**	

PM Instance and Data Planning

The following describes how to configure PM on a fiber to the mobile base station (FTTM) network.

Figure 31-35 LM on an FTTM network



Assume that mobile services in Figure 31-35 are transmitted upstream through the SPUF board on the OLT. If maintenance engineers need to perform an LM test on NodeB 1's voice services on an FTTM network, deploy the MEPs on the MDU user port and the OLT upstream port as shown in Figure 31-35. The MEP on the OLT functions as the originating end, and the MEP on the MDU functions as the responding end. The following table lists the data planning.

Item	Data
MEP	The MEP ID and the port where the MEP is configured are shown in Figure 31-35. MEP level: 0 VLAN planning of the MEP: N:1 VLAN (as listed in the following) • Customer VLANs (C-VLANs) are used to differentiate between upstream service flows of a NodeB. • The MDU translates C-VLANs to service VLANs (S-VLANs). Packets of the same type are transmitted upstream through the same S-VLAN. S-VLAN 100 is intended for data services, while

Item	Data
	S-VLAN 200 is intended for voice services. • The OLT transparently transmits VLANs.
PM instance	Test ID planning: • Data services: test ID 100 • Voice services: test ID 200 QoS planning: The MDU and OLT copy the priority of user-side packets as the priority of upstream packets, and copy the priority of network-side packets as the priority of downstream packets. • Data services: The priority of upstream packets is 0 and that of downstream packets is 1. • Voice services: The priority of both upstream and downstream packets is 5. Onward MAC and backward MAC: The device randomly chooses the source MAC addresses of the service packets sent from a NodeB as the onward and backward MAC addresses.
PM task	The Ethernet protocol packets used for an LM test are sent every 1s and for 4 times.

Procedure

Configure MEPs.

MDU configurations:

```
MDU(config)#cfm md 0 name-format no-name level 0
MDU(config)#cfm ma 0/1 name-format icc-based FTTM_VOICE
MDU(config)#cfm ma 0/1 vlan 200
MDU(config)#cfm ma 0/1 meplist 1
MDU(config)#cfm ma 0/1 meplist 2
MDU(config)#cfm mep 0/1/2 direction up port 0/2/0 vlantag1 200
```

OLT configurations:

```
OLT(config)#cfm md 0 name-format no-name level 0
OLT(config)#cfm ma 0/1 name-format icc-based FTTM_VOICE
OLT(config)#cfm ma 0/1 vlan 200
OLT(config)#cfm ma 0/1 meplist 1
OLT(config)#cfm ma 0/1 meplist 2
OLT(config)#cfm mep 0/1/1 direction up port 0/1/0 vlantag1 200
```

Step 1 Configure PM instances.

MDU configurations:

```
MDU(config)#pm instance test-id 200 mep 0/1/2 remote-mep 1 receive-8021p 5 send-8021p
5 update-backward-mac
```

OLT configurations:

```
OLT(config)#pm instance test-id 200 mep 0/1/1 remote-mep 2 receive-8021p 5 send-8021p
5 update-onward-mac
```

Step 2 Start a PM task and obtain the test results.

MDU configurations:

```
MDU(config)#pm lm receive single-ended test-id 200
```

OLT configurations:

```
OLT(config)#pm lm send single-ended test-id 200 interval 1000 count 4

 Command:
       pm lm send single-ended test-id 100 interval 1000 count 4
 Press CTRL_C to break
 ---------------------------------------------------------------------
 Index   Near-loss   Near-loss    Far-loss    Far-loss   ERR
                  ratio                  ratio
 ---------------------------------------------------------------------
    1          0          0          8     7407407     1
    2          0          0        100     1000000     0
    3       8000    80000000        900    90000000     0
 ---------------------------------------------------------------------
 Maximum near-loss : 8000       Maximum near-loss ratio : 80000000
 Minimum near-loss : 0          Minimum near-loss ratio : 0
 Average near-loss : 2666       Average near-loss ratio : -
 Total near-loss   : -
 Maximum far-loss  : 900        Maximum far-loss ratio  : 90000000
 Minimum far-loss  : 8          Minimum far-loss ratio  : 1000000
 Average far-loss  : 336        Average far-loss ratio  : 9074540
 Total far-loss    : 1008
 Note:
    Ratio unit: 10^(-8)
    ERR:
      0: No error
      1: Statistics for single-ended LM are incorrect
      2: Statistics for average single-ended LM ratio are incorrect
      3: Statistics for two-way DM are incorrect
      4: A value exceeds the threshold and therefore a reverse occurs
```

Step 3 Stop a PM task.

In this test, the PM task is completed on the originating end. Maintenance engineers only need to run the following command to stop the PM task on the responding end. MDU configurations:

```
MDU(config)#undo pm lm receive single-ended test-id 200
```

----End

Result

The LM test result (including the number of lost packets on the local and remote ends and packet loss rate) will be displayed on the CLI, as listed in step 3. ITU-T Y.1731 provides the detailed explanation about the parameters in the test results.

Follow-up Procedure

Maintenance engineers diagnose whether the link between the 2 MEPs is faulty based on the test results.

32 Clock and Time Feature

About This Chapter

This topic describes the feature of the clock and time system.

32.1 Network Synchronization Requirements

Definition

IP-nization is the trend of future network and service development, so is the trend of the bearer network. Difficulties, however, currently exist in the transition from the SDH-based traditional network to the IP-based Ethernet bearer network. One key technology involved is how to bear traditional TDM service on the new network. Traditional TDM service has two major applications: voice service and clock synchronization service.

In a traditional communications network architecture, the TDM service of the fixed network is mainly voice service. Cumulative inconsistency between the clocks at both ends of the bearer network over a long time causes bit slip. The ITU-T Recommendation G.823 defines the requirements on and the test standards of the TDM service of the fixed network. The definition is called the G.823 traffic interface standard. Apart from the bearer network, a traditional communications network usually contains an independent clock-issuing network, which adopts PDH/SDH for issuing clock signals. As specified by the ITU-T, the clock must meet the G.823 TIMING interface requirements.

In a communications network, the wireless application has the most rigorous requirements on the clock frequency. The frequencies of different BTSs must be synchronized within a specified precision. Otherwise, re-synchronization occurs during the BTS switching. Current wireless technologies are in different systems. Different systems have different requirements on the clock bearing. European systems, of which the GSM/WCDMA is a representative, adopt the asynchronous base station technologies. In this case, only frequency synchronization is required, at a precision of 0.05 ppm (or 50 ppb). The clock needs to be provided by the bearer network. The traditional solution is to provide the clock through PDH/SDH. After the IP-nization, the clock needs to be provided by the IP network. The synchronous BTS technologies, of which the CDMA/CDMA2000 is a representative, require phase synchronization of the clock (also called time synchronization). Table 32-1 lists the detailed requirements on clocks.

NOTE
Clock synchronization is frequency synchronization.

Time synchronization is phase synchronization, which requires both phase synchronization and frequency synchronization at the same time.

Table 32-1 Requirements of different standards on the clock and time

Wireless System	Frequency Precision	Phase Synchronization Precision
GSM	0.05 ppm	N/A
WCDMA	0.05 ppm	N/A
TD-SCDMA	0.05 ppm	±1.5 us
CDMA2000	0.05 ppm	±3 us
WiMax FDD	0.05 ppm	N/A
WiMax TDD	0.05 ppm	±0.5 us
LTE TDD	0.05 ppm	±1.5 us

Purpose

The purpose is to ensure the clock synchronization between communications devices and communications networks.

32.2 Synchronization Overview

Generally, two synchronization methods are used: frequency synchronization and phase synchronization. Frequency synchronization is named clock synchronization, and phase synchronization is named time synchronization.

The synchronization methods are used depending on system types. For SDH networks, frequency synchronization is used, requiring that NEs on the network synchronize their frequencies to ensure normal transmission of SDH services. Wireless stations form a time synchronization system, requiring time synchronization between neighboring stations to ensure switching among stations.

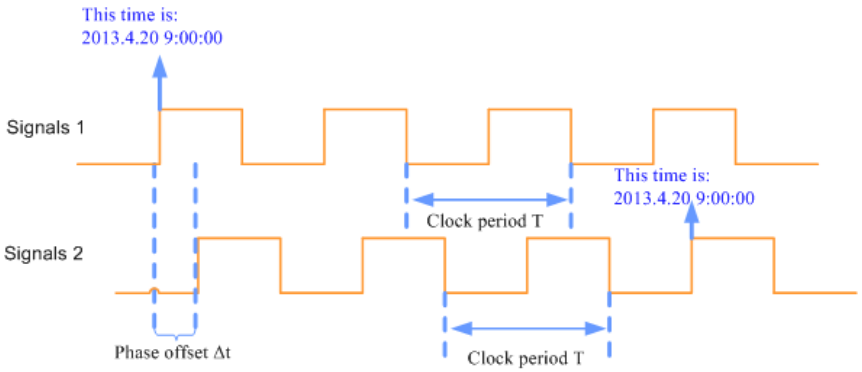
Frequency Synchronization

Frequency synchronization, also called clock synchronization, means that signals maintain a specific relationship in frequency or phase. Significant instants of a clock signal occur at the same rate to ensure that equipment in the entire communication network runs at the same rate. In other words, signals maintain a constant phase offset.

Figure 32-1 shows two clocks that synchronize their frequencies to each other. The relationship between the two clocks is summarized as follows:

- The two clocks have the same clock frequency (their clock periods Ts are the same).
- The clock pulses of the two clocks may have different phases (the phase difference is not 0).

Figure 32-1 Frequency synchronization



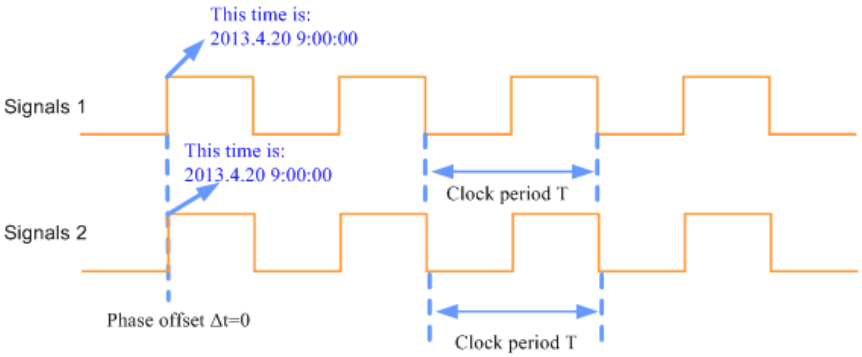
Phase Synchronization

Phase synchronization means that different signals have the same frequency and the same start and end pulse time. It is independent of the pulse sequences. Phase synchronization, also called time synchronization, means that signals have the same frequency and phase. That is, there is no phase offset between signals.

Figure 32-2 shows two clocks providing phase synchronization. The relationship between the two clocks is summarized as follows:

- The two clocks have the same clock frequency (their clock periods Ts are the same).
- Clock pulses of the two clocks have the same phase (the phase difference is 0).

Figure 32-2 Phase synchronization



32.3 Clock Synchronization

The following figure shows the clock synchronization solutions supported by the MA5600T/MA5603T/MA5608T.

Figure 32-3 Clock synchronization solutions

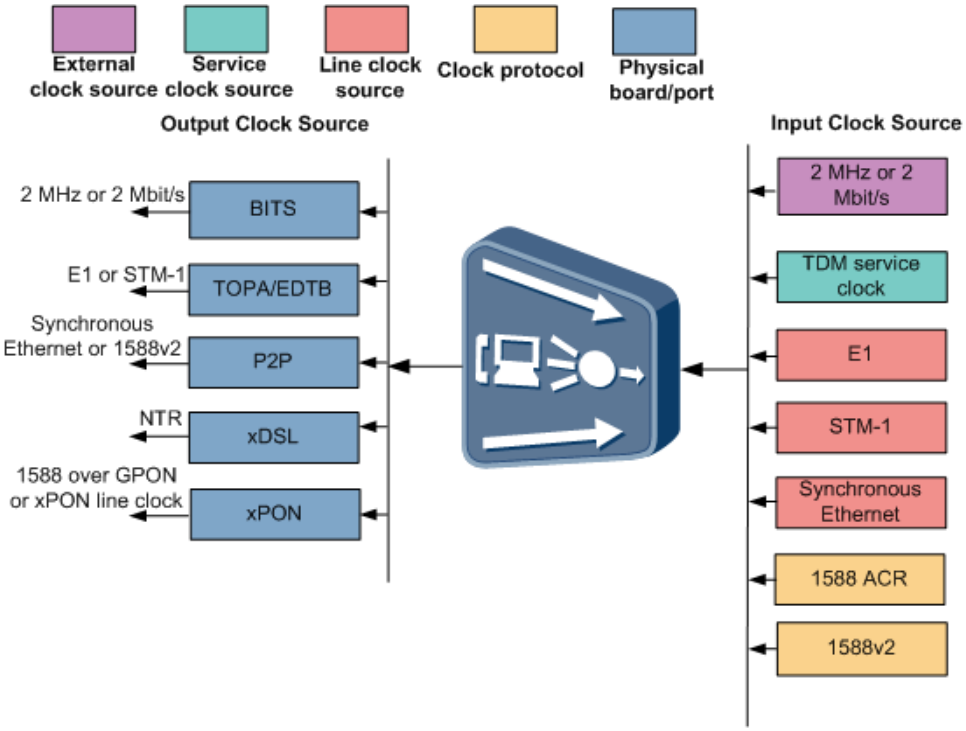


Table 32-2 Comparison between clock synchronization solutions

Solution	Description	Remarks
2 MHz or 2 Mbit/s BITS clock	The MA5600T/MA5603T/MA5608T traces a BITS input clock source and supports 2.048 Mbit/s bitstream (HDB3) or 2.048 MHz clock signal input.	This solution features high deployment costs because each CO requires a BITS device.
TDM service clock	The MA5600T/MA5603T/MA5608T traces the adaptive clock source recovered from TDM service bitstreams.	This solution promotes high quality requirements on MAN networks. Huawei does not recommend this solution for carriers.
E1 or STM-1	The MA5600T/MA5603T/MA5608T traces the clock source recovered from E1 or STM-1 lines. The MA5600T/MA5603T/MA5608T can also use the clock recovered from line bitstreams as the system clock source.	This solution applies in traditional TDM private line service scenarios.
Synchronous Ethernet	Synchronous Ethernet uses Ethernet bitstreams to recover clock signals for Ethernet clock synchronization. The implementation mode is similar to that on an SDH or PDH network. The MA5600T/MA5603T/MA5608T uses	This solution requires bearer networks to support synchronous Ethernet.

Solution	Description	Remarks
	high-precision clock signals as the transit reference in the transmit direction. It recovers and extracts the clock signals at the receive end. The physical layer transmits and receives the clock signals and is compatible with traditional networks. The MA5600T/MA5603T/MA5608T traces the clock source recovered from GE or 10GE lines. The MA5600T/MA5603T/MA5608T can also use the clock recovered from line bitstreams as the system clock source.	
1588v2	The IEEE organization proposed the IEEE 1588v2 precision time synchronization protocol, which supports system-wide synchronization accuracy within the sub-microsecond range. In contrast to the GPS solution, the IEEE 1588v2 solution achieves the same time precision but has advantages in terms of cost, maintenance, and security. It has become the most popular time synchronization protocol in the industry.	This solution requires bearer networks to support 1588v2 at each hop and applies in mobile bearing scenarios that require time synchronization, such as CDMA2000, TD-SCDMA, and LTE.
1588 adaptive clock recovery (ACR)	1588 ACR synchronizes the frequencies of the devices at the two ends of a packet switched network (PSN). Specifically, a master device supporting 1588v2 encapsulates the local system clock data into 1588v2 packets. The PSN network transparently transmits the 1588v2 packets to a slave device. The slave device obtains timestamps from the received 1588v2 packets and recovers the clock data of the master device. The PSN network does not need to support clock synchronization, or the clock of the PSN network can be a third-party clock.	• Provides E2E clock synchronization and is simple in deployment, meeting clock frequency requirements of MSAN and wireless networks. • Promotes packet delay variation (PDV) requirements on intermediate networks.
Network Timing Reference (NTR)	NTR synchronizes transmit clock frequencies of xDSL lines at the physical layer. The functions of NTR are similar to those of synchronous Ethernet. With either of the functions enabled, the MA5600T/MA5603T/MA5608T transmits clock signals and recovers a clock from serial bitstreams of xDSL lines for clock synchronization.	This solution applies only in DSL lines and requires terminals to support NTR.
PON line clock	The MA5600T/MA5603T/MA5608T traces the clock source recovered from PON lines. The MA5600T/MA5603T/MA5608T can also use the clock recovered from line bitstreams as the	None

Solution	Description	Remarks
	system clock source.	

Table 32-3 Comparison between clock synchronization solutions—Hardware Requirements

Solution	OLT	MA5616	MA5694/MA5698/MA5898	MA5811S
2 MHz or 2 Mbit/s BITS clock	Yes	Yes	Yes	Yes
TDM service clock	Yes	Yes	MA5698/MA5898: Yes MA5694: N/A	Yes
E1 or STM-1	Yes	Yes	MA5698/MA5898: Yes MA5694: N/A	N/A
Synchronous Ethernet	Yes	Yes	Yes	Yes
1588v2	• Clock Input: Yes • Clock Output – 1588v2 over PON – 1588v2 over GE	No	MA5698/MA5898: Yes MA5694 (PON upstream transmission): Yes MA5694 (PON/VDSL2 bonding upstream transmission): No	No
1588 adaptive clock recovery (ACR)	Yes	No	No	No
Network Timing Reference (NTR)	Yes	No	No	Yes
PON line clock	Yes	Yes	Yes	Yes

32.4 Time Synchronization

Figure 32-4 Time synchronization solutions

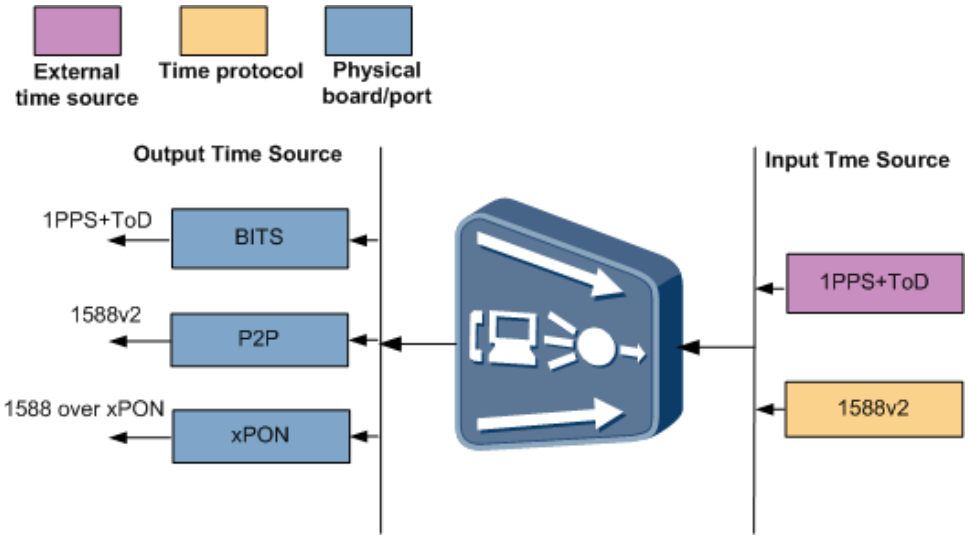


Table 32-4 Comparison between time synchronization solutions

Solution	Description	Remarks	Time Input	Time Output
1 PPS+ TOD	The MA5600T/MA5603T/MA5608T traces the input time source and supports 1 PPS+TOD time synchronization.	In some mobile bearing scenarios, some intermediate devices do not support time transfer. Therefore, the 1 PPS+TOD time must be sent to the MA5600T/MA5603T/MA5608T and the MA5600T/MA5603T/MA5608T uses the 1 PPS+TOD time as the system time source. This solution features high deployment costs because each CO requires BITS and GPS devices.	Yes	Yes
1588v2	The IEEE organization proposed the IEEE 1588v2 precision time synchronization protocol, which supports system-wide synchronization accuracy within the sub-microsecond range. In contrast to the GPS solution, the IEEE 1588v2 solution	This solution requires bearer networks to support 1588v2 at each hop and applies in mobile bearing scenarios that require time synchronization, such as CDMA2000, TD-SCDMA, and LTE.	Yes	• 1588v2 over PON • 1588v2 over GE

Solution	Description	Remarks	Time Input	Time Output
	achieves the same time precision but has advantages in terms of cost, maintenance, and security. It has become the most popular time synchronization protocol in the industry.			

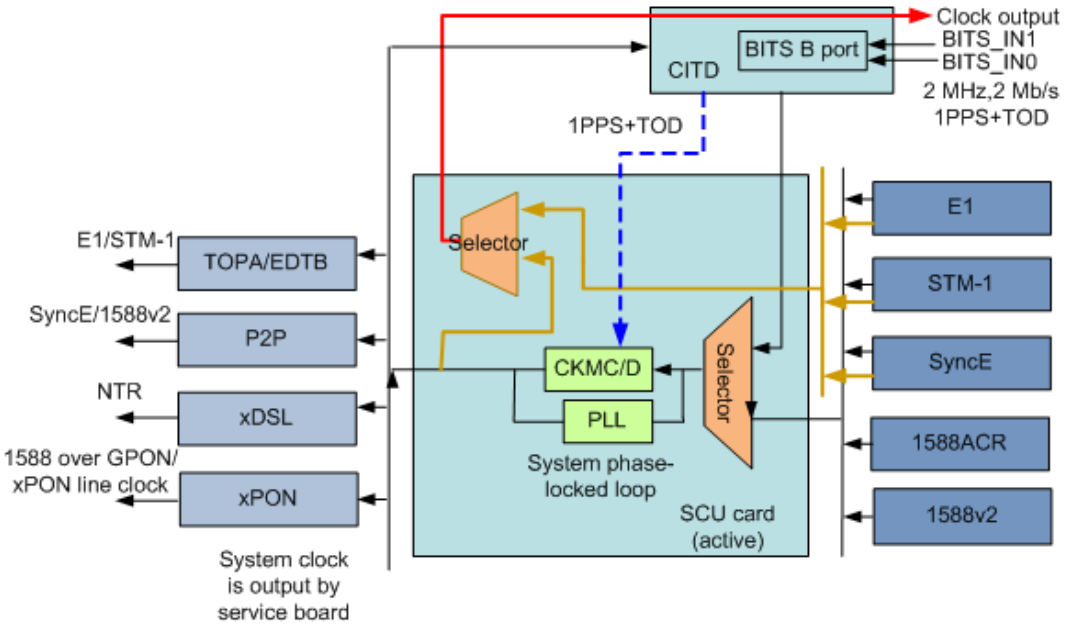
32.5 Physical Layer Clock/Time Synchronization

This section describes the definition and principle of the physical clock and time system of the MA5600T/MA5603T/MA5608T, and describes the specific applications of physical clock and time synchronization.

32.5.1 Physical Layer Clock/Time Synchronization Principles

The clock system of the MA5600T/MA5603T/MA5608T consists of three parts: system clock/time synchronization source, phase-locked loop circuits, and clock/time output, as shown in Figure 32-5.

Figure 32-5 Solution for the clock and time synchronization system



NOTE

The stratum-3 clock daughter boards of the MA5600T/MA5603T/MA5608T are integrated as the CKMC. The CKMC daughter board supports the clock performance required by G.813 and G.8262, and the active/standby configuration. As a BITS interface board, the CITDboard provides two BITS inputs and one BITS output.

When the system is not configured with a stratum-3 clock daughter board, the hardware phase-locked loop of the control board can be used to provide the system clock with a precision of ±25 ppm.

Configuring the System Phase-Locked Loop

The stratum-3 clock unit of the system is an optional unit. When configured with the CKMC clock unit, the system provides high-quality clock output that meets the requirements of G.813 and G.8262.

When the system is not configured with the stratum-3 clock unit, the system clock is provided by the hardware phase-locked loop. The quality of the clock thus provided meets the requirements of G.8261 or G.823. The quality of the clock provided for the SAToP service and native TDM service meets the requirements of G.8261 SAToP or G.823 traffic. The quality of the locked ideal line clock meets the requirements of G.8261 EEC or G.823 synchronization.

In the case of the SAToP service and native TDM service, the clock quality is not directly related to the stratum-3 clock unit. The clock quality meets the requirements of G.8261 SAToP or G.823 traffic.

Clock/Time Output

The MA5600T/MA5603T/MA5608T system outputs clock signals through the clock output interface or the synchronization service interface. The output clock signals serve as the reference clock of the device interconnected with the MA5600T/MA5603T/MA5608T. The synchronization service interface can select the system clock, line receive clock, board oscillator, or SAToP clock as the transmit clock. The interface is capable of outputting the system clock signals only when the system clock is selected as the transmit clock of the interface.

The MA5600T/MA5603T/MA5608T system outputs clock signals through the clock output interface or the synchronization service interface. The output clock signals serve as the reference clock of the device interconnected with the MA5600T/MA5603T/MA5608T. The synchronization service interface can select the system clock, line receive clock, board oscillator, or CESoP clock as the transmit clock. The interface is capable of outputting the system clock signals only when the system clock is selected as the transmit clock of the interface.

Line Clock Output

The synchronization service ports of the MA5600T/MA5603T/MA5608T include the STM-1, PDH, synchronous Ethernet, xDSL, and xPON ports.

The STM-1 port can select the system clock or the line receive clock of the port itself as the line transmit clock.

The PDH port is mainly used for the E1 mode (hardware compatible with the T1 mode). In the E1 mode, the system clock or the line clock can be selected as the transmit clock.

The transmit clock of the synchronous Ethernet port is the system clock by default. The MA5600T/MA5603T/MA5608T supports the change of the transmit clock mode of the synchronous Ethernet port. The Ethernet port without the synchronization capability adopts the oscillator of the board as the transmit clock.

The MA5600T/MA5603T/MA5608T functions as an OLT, and the line transmit clock of the PON port is used as the system clock, the signals of which are transmitted to the ONT.

The xDSL port with the synchronization capability is mainly used for the G.SHDSL emulation service. In the application of the emulation service, the G.SHDSL NTR clock should be set as the system clock.

Clock Output

The MA5600T/MA5603T/MA5608T provides one output clock, which is configurable. The system clock or the line clock can serve as the output clock. When the line clock is used, the system directly exports the signals of the line clock source and outputs the signals through the clock output interface. When the system clock is used, the system phase-locks the clock signals and outputs the signals through the clock output interface.

The output interface board of the MA5600T or MA5603T is CITD.

The output board of the MA5608T is H801MCUD1.

Tributary Clock Output

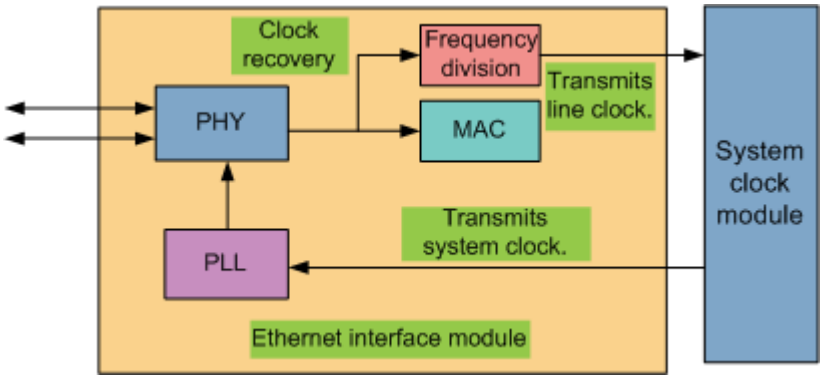
The TOPA board (configured with the CSSA daughter board) is used for SAToP based on STM-1. The system clock, tributary receive clock, or SAToP recovered clock can be selected as the tributary E1 transmit clock of the TOPA board (configured with CSSA).

- System clock: The tributary E1 transmit clock synchronizes with the system clock signals issued by the control board.
- Tributary receive clock: The tributary E1 transmit clock synchronizes with the receive clock of the same port.
- SAToP clock: The tributary E1 transmit clock synchronizes with a certain SAToP recovered clock. The recovery source of the SAToP clock can be flexibly selected as long as the SAToP is recovered within the board.

Synchronous Ethernet

Synchronous Ethernet, similar to traditional SDH synchronization, is a synchronization technology that uses physical-layer bitstreams to carry frequency information and recovers frequency information also from physical-layer bitstreams. Synchronous Ethernet enables the MA5600T/MA5603T/MA5608T to extract clock signals from serial bitstreams on an Ethernet line, select a better clock source, and finally send the clock signals (line clock) to a system clock phase-locked loop (PLL). The PLL then traces the clock signals and generates a system clock. The MA5600T/MA5603T/MA5608T uses the system clock as the transmit clock at the Ethernet physical layer to transmit clock data to lower-layer devices.

Figure 32-6 Synchronous Ethernet clock signal processing



- Ethernet interface receive clock:
  - The PHY chip in the Ethernet interface module recovers a clock from Ethernet line bitstreams, divides the frequency of the clock, and sends the clock to the system clock module.
  - The system clock module selects the clock source with the highest precision based on clock source priorities and sends the clock source to the system PLL as the system reference clock. Then, the PLL traces the reference clock and outputs a high-precision clock for each interface.
- Ethernet interface transmit clock:
  - The system clock module sends a high-precision clock to Ethernet interfaces on NEs.
  - The PLL of the Ethernet interface module traces the high-precision clock sent from the system clock module, generates a reference clock for the Ethernet link bitstreams, and sends the reference clock out using Ethernet link bitstreams.

Performance and Deployment Limitations

A synchronous Ethernet clock is simple to implement. The mechanism for transferring the clock in synchronous Ethernet is simple to implement and the recovered clock is reliable in compliance with ITU-T G.823 with respect to the synchronization interface specifications.

Similar to an SDH network, a synchronous Ethernet network has limitations on deployment. For the synchronous Ethernet network, the clock is transferred over links. All nodes on the clock trail must have the synchronous Ethernet feature to achieve clock synchronization (only frequency synchronization) of the entire network.

Synchronous Ethernet can work with IEEE 1588v2 clocks for time synchronization.

External Time Source

A time source is a signal source containing reference timing information. Each NE synchronizes its local clock phase to the reference timing using its phase-locked loop (PLL). In this manner, time synchronization is implemented on the entire network.

A time source can be from an external time source or a line time source.

In some mobile carrying scenarios, 1PPS+TOD time signals must be injected to an OLT or an MDU because intermediate devices do not support transmission of these signals. Then the signals are used as the system time source.

The 1PPS+TOD protocol receives signals using a solution similar to time signal reception using a GPS satellite. This protocol transmits two types of signals: 1PPS and TOD signals. Frame headers and time information are transmitted over dual lines.

1PPS+TOD time signals consist of 1PPS signals and TOD time information.

- 1PPS

  1PPS is short for one pulse per second. 1PPS signals are used for time scaling and work at the RS-422 levels. The pulse frequency of 1PPS is 1 Hz. That is, one pulse is transmitted per second. The 1PPS signal pulse width ranges from 20 ms to 200 ms. On the rising edge of the pulse, UTC time signals are aligned.

- TOD

  TOD is short for time of day. TOD messages provide time in ASCII format. TOD signals also work at the RS-422 or RS-232 levels and provide a baud rate of 9600 bit/s. A TOD message contains information such as current date/time, time standard ID, 1PPS status flag, date/time adjusted based on UTC leap seconds, leap second adjustment directive, and GPS time.

ToD can be in UBX or NMEA format.

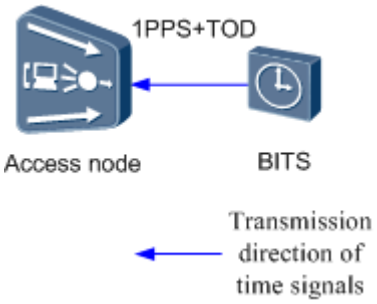
- UBX: is a private format of the Ublox company. The UBX format is in binary code and identified by a length domain because it does not contain a start or end character.
- NMEA: is a set of communication protocol formulated by the National Marine Electronics Association to standardize various global positioning system (GPS) navigators. The NMEA protocol is in ASCII format and identified by start and end characters. A comma (,) is used to separate domains.

NOTE
The OLT supports only the UBX format.

As shown in Figure 32-7, the OLT or MDU can function as an access node to receive 1PPS+TOD clock signals.

- When the OLT functions as the access node, it receives 1PPS+TOD clock signals using the H801CITD/H806VPEA board as the clock source.
- When the MDU functions as the access node, it receives 1PPS+TOD clock signals using the clock/time output port CLK/TOD0 port as the clock source.

Figure 32-7 External time source synchronization



Clock Mode

This topic describes the three clock modes supported by the Router.

The Router supports the following clock modes: Locked mode, Holdover mode and Free-run mode.

Locked Mode

In this mode, the system clock source is synchronized with the input clock source. The phases of the system clock source and the input clock source are in a constant relation. The Router traces the BITS clock source, the upstream Ethernet synchronous clock, line clock source, Auto-adaptation recovery clock , or the IEEE1588 V2 message recovered clock source.

Holdover Mode

The MA5600T/MA5603T/MA5608T records the clock data of the trace mode. If the clock source is lost, the system builds a system clock by using the recorded clock data, and maintains the clock properties as consistent as possible with the clock properties in the trace mode. As such, the system enters the holdover mode. The precision of the holdover mode meets the G.813 or G.8262 requirements.

The system supports the free-run mode only when configured with the stratum-3 clock daughter board.

Free-Run Mode

In this mode, the Router works based on the inherent frequency of the internal crystal oscillator inside the device.

Clock Source Selection

One MA5600T/MA5603T/MA5608T supports a maximum of 10 external clock sources. It uses a clock source selection algorithm to select a clock source as its reference clock. The MA5600T/MA5603T/MA5608T supports the clock source selection algorithm based on clock source priorities or synchronization status messages (SSMs).

Clock Source Priority-based Clock Source Selection

If the MA5600T/MA5603T/MA5608T uses the clock source priority-based clock source selection algorithm to select a clock source, users need to configure the external clock sources with different priorities. The MA5600T/MA5603T/MA5608T then automatically selects the clock source that has the highest priority and is running properly as its reference clock. When the traced clock source becomes faulty, the MA5600T/MA5603T/MA5608T automatically selects the clock source that has the second highest priority and is running properly as its reference clock.

SSM-based Clock Source Selection

SSM messages are used for transporting clock signal quality levels over clock links. By reading data carried in the SSM messages, the node clocks on SDH and synchronization networks obtain the information of the upper-layer clocks, perform operations based on instructions carried in the SSM messages, such as tracing a clock, switching a clock source, or entering the holdover state, and send the clock synchronization information of the node clocks to lower-layer clocks.

The following table lists ITU-T G.781-compliant SSM quality levels.

Quality Level	SSM Code	Code Value	Description
QL-PRC	0010	0x02	Indicates that the timing quality generated by a primary reference clock defined in Recommendation G.811 is transported.
QL-SSU-A	0100	0x04	Indicates that the timing quality generated by a type I or V (transit exchange) slave clock defined in Recommendation G.812 is transported.
QL-SSU-B	1000	0x08	Indicates that the timing quality generated by a type VI (local exchange) slave clock defined in Recommendation G.812 is transported.
QL-SEC	1011	0x0b	Indicates that the timing quality generated by an SDH equipment clock (SEC) defined in Recommendation G.813 is transported.
QL-DNU	1111	0x0f	Indicates that the clock source is not used for synchronization.

If the SSM-based clock source selection algorithm is used, the MA5600T/MA5603T/MA5608T extracts the quality level from clock source input signals. If the clock source does not support SSM information extraction, users must manually configure the quality level. The MA5600T/MA5603T/MA5608T automatically selects the clock source that has the highest quality level and is running properly as its reference clock. Ensure that the quality level is higher than or equal to the lowest SSM level of the clock source. If there are two external clock sources with the same highest quality level, the MA5600T/MA5603T/MA5608T selects the one that has a higher clock source priority. When the clock source traced by the MA5600T/MA5603T/MA5608T becomes faulty, the MA5600T/MA5603T/MA5608T automatically selects the clock source that has the second highest quality level and is running properly as its reference clock.

In SSM-based clock source selection mode, SSM quality levels of clock sources are used for selecting a clock source. The MA5600T/MA5603T/MA5608T obtains the SSM quality levels of clock sources in either of the following modes:

- Users configure the SSM quality levels of clock sources. In this mode, all clock source ports support SSM quality level configurations. For the clock source ports that support the extraction of SSM quality levels, users can also configure SSM quality levels for the clock source ports as the SSM quality levels of these clock sources.
- The MA5600T/MA5603T/MA5608T extracts SSM information, such as the S1 byte in STM-1, from lines. Not all boards support the extraction of SSM information from lines. For details, see section "Feature Dependencies and Limitations."

The MA5600T/MA5603T/MA5608T uses Ethernet (SSM), E1, STM-N, or PON ports to transport clock quality levels to lower-layer devices.

- STM-N ports use the lower 4 bits of the section overhead byte S1 to transport quality levels.

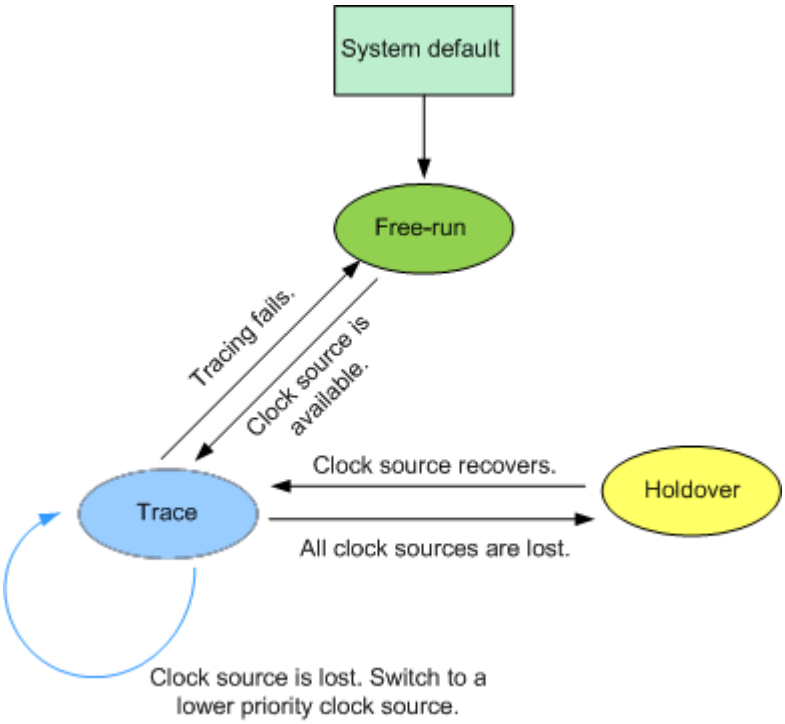
- E1 (2 Mbit/s) ports use four idle bits in the TS0 timeslot to transport quality levels.
- Ethernet clock synchronization ports use ESMC packets to transport quality levels.

The MA5600T/MA5603T/MA5608T enters holdover if no clock source is available, regardless of whether the clock source selection algorithm is based on clock source priorities or SSM messages. In holdover mode, the MA5600T/MA5603T/MA5608T uses the data recorded in tracing mode to control the clock unit so that the clock unit generates clock properties similar to those of the clock in tracing mode.

- If clock recovery is enabled, the MA5600T/MA5603T/MA5608T enters the tracing mode from the holdover mode when a clock source recovers.
- If clock recovery is disabled, the MA5600T/MA5603T/MA5608T does not enter the tracing mode, regardless of whether a clock source recovers.
- If the MA5600T/MA5603T/MA5608T restarts during the holdover duration, the system enters the free-run mode

Figure 32-8 shows the clock source status transition.

Figure 32-8 Clock source status transition



32.5.2 Physical Layer Clock/Time Synchronization Usage Scenarios

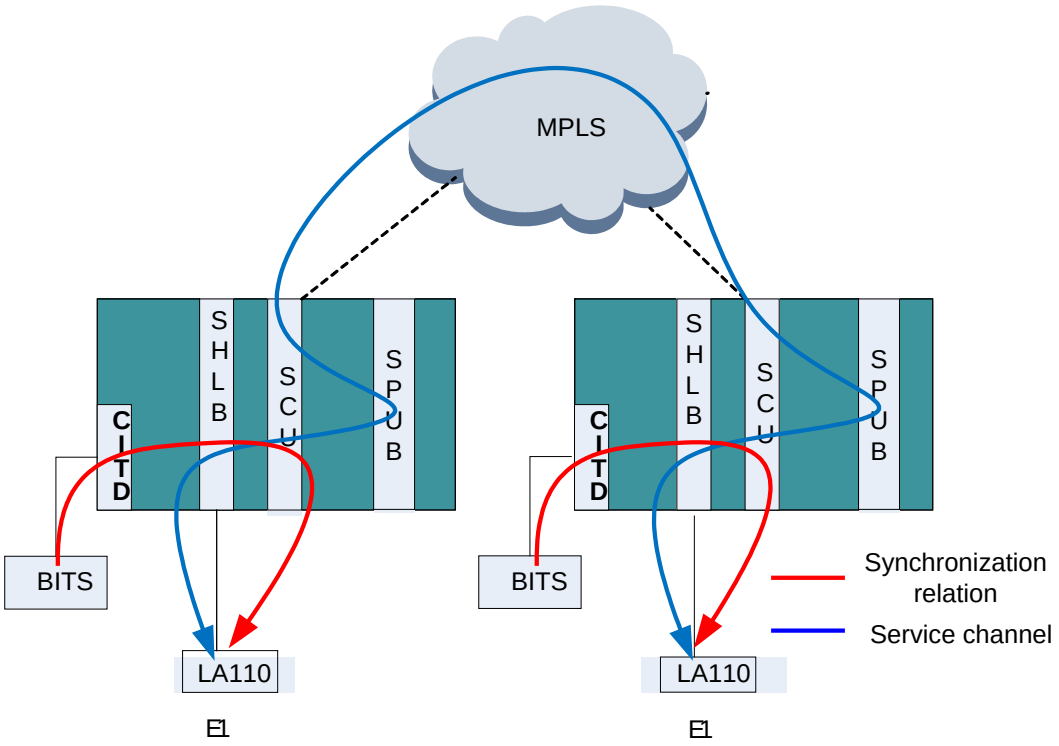
Clock Synchronization of the SHDSL PWE3 Service

As a DSLAM, the MA5600T/MA5603T/MA5608T is mainly used in the xDSL access scenario, which usually does not require synchronization of the system clock. In the SHDSL access, however, the synchronization property of the SAToP emulation service needs to be considered. For example, in the PWE3 service and the ATM emulation service, the RTU needs

to synchronize with the CO device and obtain the NTR network clock signals. Therefore, when configuring the SHDSL board, set the NTR clock, instead of the oscillator of the SHDSL board, as the system clock. This is to implement the end-to-end RTU synchronization.

Figure 32-9 illustrates the PWE3 networking of the MA5600T/MA5603T/MA5608T through G.SHDSL. The networking implements the end-to-end private line service between E1 ports. Transmission is implemented over the service channel in the packet network, and a direct synchronization relation is not involved. The two MA5600T/MA5603T/MA5608Ts are synchronized with the BITS clock (the MA5600T/MA5603T/MA5608Ts can also be synchronized in other ways). The two SHLB boards respectively issue the system clock signals (NTR) of the MA5600T/MA5603T/MA5608Ts to the LA110 terminals. In this way, the two terminals are synchronized with the network clock.

Figure 32-9 Clock synchronization of the SHDSL PWE3 service



Hardware configuration of the system:

- Board configuration: SHLB, SCU, and SPUB.
- Clock configuration: The stratum-3 clock unit is an optional configuration. When the stratum-3 clock unit is configured, the clock quality meets the G.813 requirements; when the stratum-3 clock unit is not configured, the clock quality meets the G.823 traffic requirements.

Key points of the synchronization configuration:

- Make sure that the MA5600T/MA5603T/MA5608Ts are synchronized, which can be implemented through the BITS, PDH, SDH, and synchronous Ethernet ports.
- When the SHDSL board is used for the SAToP service, select the line clock as the system clock of the NE.

- Set the locked network clock of the LA110 terminal, and issue the setting to the E1 port, thus realizing end-to-end synchronization between E1 ports.
- The device connected to the E1 port locks the E1 line clock of the LA110 terminal, thus realizing clock synchronization of the entire network.

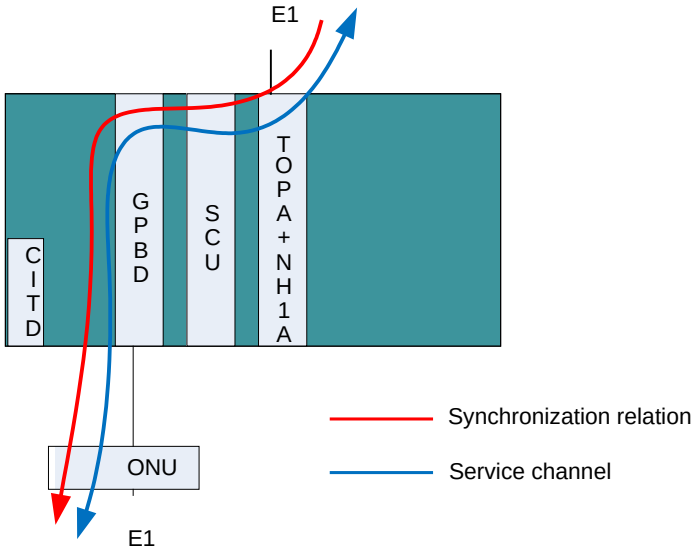
Clock Synchronization of the Native TDM Service

The MA5600T/MA5603T/MA5608T can carry native TDM service through GPON and provide multi-E1 end-to-end service. The E1s can be asynchronous with each other. The E1 signals go upstream through the GPON terminal, then undergo the VC12 adaptation and GEM encapsulation, and are then switched to the TOPA board through the GPBD board and the control board. Then, the TOPA board performs GEM decapsulation and VC12 deadaptation, and implements the E1 upstream transmission or STM-1 upstream transmission through different daughter boards (16 E1 upstream transmission through the NH1A daughter board, and 2 STM-1 upstream transmission through the O2CE daughter board).

Clock Mode 1: System Clock

In this mode, the MA5600T/MA5603T/MA5608T locks the upstream E1 line clock of the TOPA board as the system clock. The system clock signals are issued to the ONU (such as the OT928G, MA5612) through the optical channel provided by the GPON port of the GPBD board. The E1 transmit clock of the ONU synchronizes with the GPON port recovery line clock. Thus, end-to-end synchronization is implemented between the upstream E1 of the TOPA board and the downstream E1 of the ONU, as shown in Figure 32-10.

Figure 32-10 System locking the E1 line clock of the TOPA board



Hardware configuration of the system:

- Board configuration: GPBD, SCU, and TOPA (configured with the NH1A daughter board).
- Clock configuration: The stratum-3 clock unit is an optional configuration. When the stratum-3 clock unit is configured, the clock quality meets the G.813 requirements; when the stratum-3 clock unit is not configured, the clock quality meets the G.823 traffic requirements.

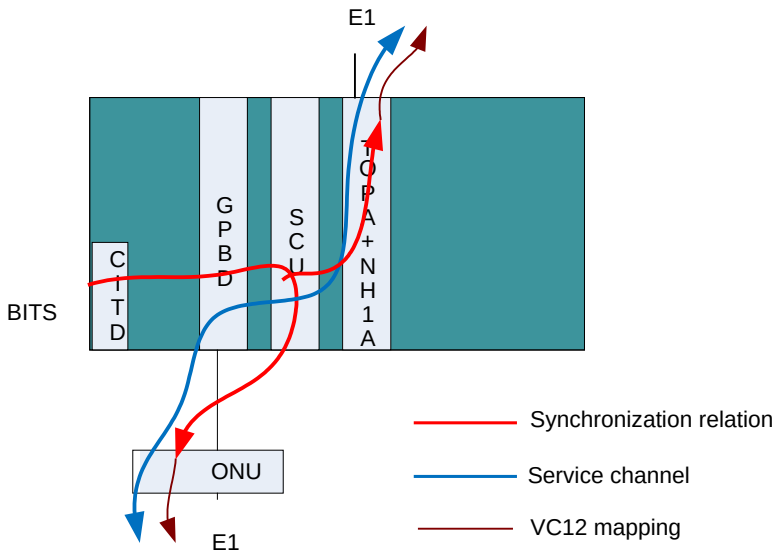
Key points of the synchronization configuration:

- The MA5600T/MA5603T/MA5608T locks an E1 line recovered clock of the TOPA board as the system clock. The device interconnected with the E1 port of the TOPA board is required to serve as the master device, and the E1 port of the TOPA board as the slave device.
- The transmit clocks of the E1s of the TOPA board need not be configured. They are the VC12 recovery clock by default, and synchronize with the E1 receive clock of the ONU (in the upstream direction).
- The transmit clock of the GPON port of the OLT is the system clock by default. The ONU synchronizes with the OLT through the PON port recovery line clock to ensure that the GEM frames are transmitted synchronously.
- The E1 transmit clock of the ONU adopts the system clock (synchronized with the OLT). It can also adopt the VC12 bit stream recovery clock, which, however, will bring unwanted VC12 adaptation jitter.
- The device connected to the ONU E1 serves as the slave device and needs to lock the transmit clock of the ONU E1, thus realizing synchronization over the entire network.

Clock Mode 2: Bit Stream Recovery Clock

In this mode, the MA5600T/MA5603T/MA5608T does not lock the upstream E1 clock of the TOPA board. That is, the system clock is asynchronous with the clock of the device interconnected with the E1 port of the TOPA board. The GPBD board issues the system clock signals to the ONU through the optical channel provided by the GPON port for synchronizing the ONU with the OLT. After the TOPA board performs the VC12 adaptation on the E1 bit stream, the signals can be transmitted to the ONU synchronously. The ONU performs the VC12 deadaptation, and recovers the clock of the downstream E1 of the TOPA board (the line receive clock of the TOPA E1). Hence, it can be seen that the E1 transmit clock of the ONU adopts the VC12 bit stream recovery clock, which is asynchronous with the system clock but synchronous with the E1 receive clock of the TOPA. In addition, all the E1s can be adapted and deadapted independently, so the E1s can be asynchronous with each other and can thus be used flexibly. During the VC12 adaptation and deadaptation of the E1s, jitter may occur, as shown in Figure 32-11.

Figure 32-11 Auto-adaptation clock recovery



Hardware configuration of the system:

- Board configuration: GPBD, SCU, and TOPA (configured with the NH1A daughter board) or TOPA (configured with the O2CE daughter board).
- Clock configuration: The stratum-3 clock unit is an optional configuration. When the stratum-3 clock unit is configured, the clock quality does not meet the G.813 requirements; when the stratum-3 clock unit is not configured, the clock quality meets the G.823 traffic requirements.

Key points of the synchronization configuration:

- The system clock of the MA5600T/MA5603T/MA5608T can be randomly configured and is usually asynchronous with the E1 line clock of the TOPA board (this is the typical application; the system clock can also be synchronous with the E1 line clock of the TOPA board).
- The transmit clocks of the E1s of the TOPA board need not be configured. They are the VC12 recovery clock by default.
- The transmit clock of the GPON port of the OLT is the system clock by default. The ONU synchronizes with the OLT through the PON port recovery line clock.
- The E1 transmit clock of the ONU adopts the VC12 bit stream recovery clock.
- The device connected to the E1 of the ONU needs to lock the E1 transmit clock of the ONU.

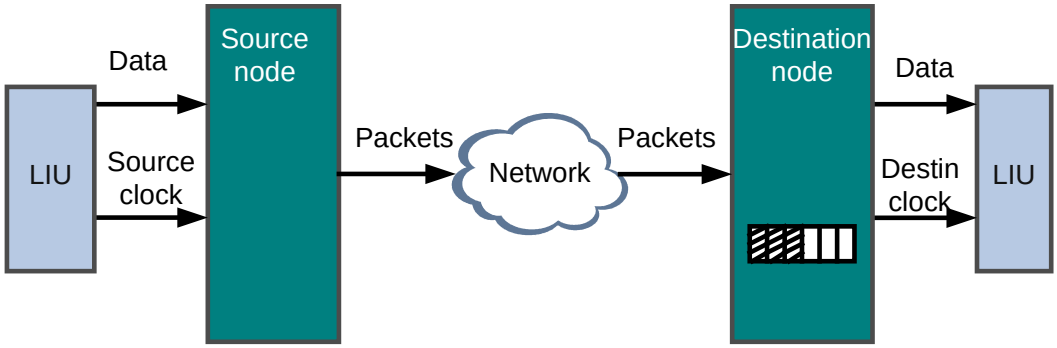
SAToP Clock Synchronization

Auto-adaptation recovery mode

Figure 32-12 illustrates the principle of the auto-adaptation clock recovery. In this mode, the receive end recovers the clock of the transmit end through the average arrival rate of the received SAToP packets. As shown in the figure, according to its source clock, the LIU (Line Unit) transmits packets to the destination device. The destination device buffers these packets in a queue, and then sends these packets according to its local clock. A least disparity between

the source clock and the local clock of the destination device will cause the depth of the buffer queue of the destination device to change. The depth of the queue can be used to determine whether the local clock is synchronous with the source clock. If the depth of the queue continuously increases, the local clock is behind the source clock. In this case, puts the local clock ahead. If the depth of the queue continuously decreases, the local clock is ahead of the source clock. In this case, puts back the local clock. The purpose of such adjustment is to ensure that the local clock is synchronous with the source clock in the long term.

Figure 32-12 Auto-adaptation clock recovery



The difficulty point of the auto-adaptation algorithm is that the IP network has inherent delay jitter, the packet delay variation (PDV). PDV can also cause change to the depth of the buffer queue. The destination LIU, however, cannot distinguish whether the change is due to the disparity of frequency or the delay jitter of the IP network, and cannot make a correct reaction. The delay jitter of the IP network is not cumulative and thus can be filtered by certain statistical methods, such as by calculating an average value.

Application of Clock Recovery in the SAToP Service

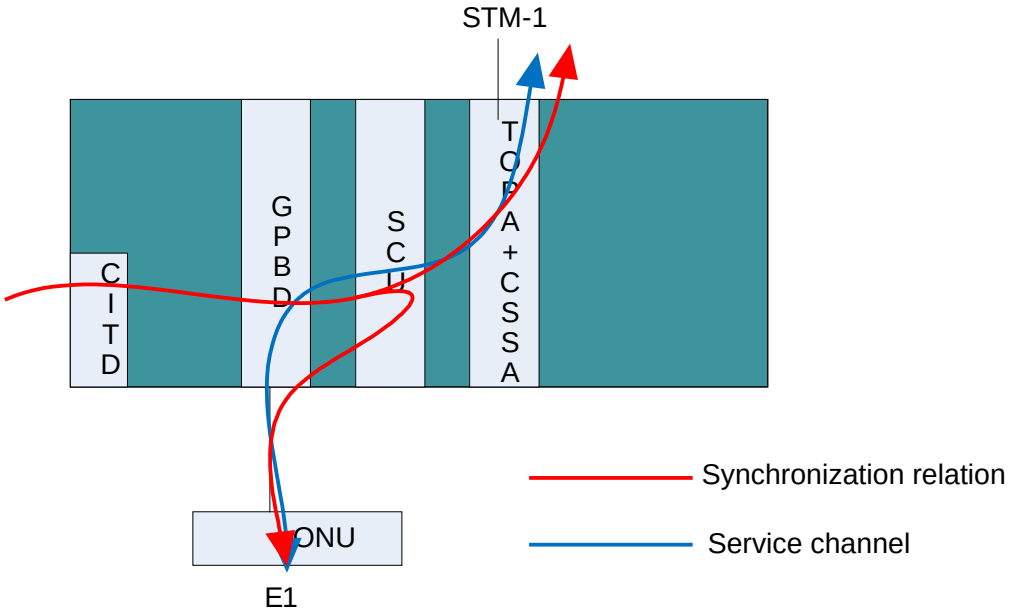
The MA5600T/MA5603T/MA5608T supports the SAToP application, can recover the clock from the SAToP data streams, and provides the end-to-end E1service. The E1 access is implemented by the GPON ONU. The E1 data is encapsulated as SAToP data, transmitted to the GPBD board, to the control board, and then goes upstream through the TOPAboard. The TOPA can implement E1 upstream transmission or STM-1 upstream transmission through different daughter boards (16 E1 upstream transmission through the EH1A daughter board; 2 STM-1 upstream transmission through the CSSA daughter board).

The transmit clock mode of each E1 port of the TOPA board is set independently. Different E1 ports can work in different clock modes. Configure the clock mode of the E1 ports according to actual application.

Clock mode 1: system clock

The E1 port of the TOPA board adopts the system clock as the transmit clock. The line transmit clock of the PON port of the GPBD board is the system clock by default. Therefore, the ONU synchronizes with the system clock of the MA5600T/MA5603T/MA5608T through the PON port, and the E1 port of the ONU adopts the line recovered clock of the PON port as the transmit clock. In this way, global synchronization is implemented from the TOPA E1 to the ONU E1, as shown in Figure 32-13.

Figure 32-13 Global synchronization through the system clock



Hardware configuration of the system:

- Board configuration: GPBD, SCU, TOPA (configured with the EH1A daughter board) or TOPA (configured with the CSSA daughter board).
- Clock configuration: The stratum-3 clock unit is an optional configuration. When the stratum-3 clock unit is configured, the clock quality meets the G.813 requirements; when the stratum-3 clock unit is not configured, the clock quality meets the G.823 traffic requirements.

Key points of the synchronization configuration:

- Configure the system clock source of the MA5600T/MA5603T/MA5608T. For example, select the BITS clock, the STM-1 line clock, the synchronization Ethernet line clock, or the oscillator as the system clock.
- Select the system clock as the transmit clock of the E1 ports of the TOPA board. For the STM-1 port, select the tributary E1 transmit clock as the system clock. Therefore, the device interconnected with the E1 port of the TOPA board is required to serve as the slave device, locking the E1 transmit clock of the TOPA board.
- The transmit clock of the GPON port of the OLT is the system clock by default. The ONU synchronizes with the OLT through the PON port recovered line clock.
- The E1 transmit clock of the ONU synchronizes with the PON port line clock, thus synchronizing with the system clock of the MA5600T/MA5603T/MA5608T.
- The device connected to the E1 port of the ONU serves as the slave device, locking the E1 transmit clock of the ONU.

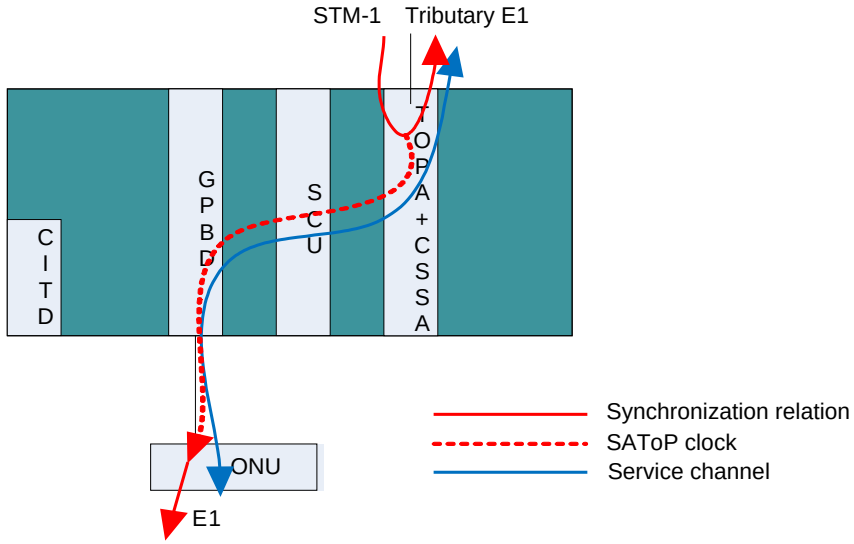
Clock mode 2: line clock

The line clock mode refers to that the E1 line receive clock is adopted as the E1 transmit clock of the TOPA board. In the networking, the port interconnected with the E1 port of the TOPA board is required to serve as the master port. The E1 data received by the upstream port is encapsulated as the SAToP data and transmitted to the ONU through the PON port. The

ONU recovers the E1 receive clock of the TOPA board through CESoP, and uses the recovered clock as the E1 transmit clock of the ONU.

When the ONU recovers the clock in the SAToP mode, the ONU can choose not to synchronize with the system clock of the MA5600T/MA5603T/MA5608T. As shown in Figure 32-14, the upstream port adopts the line clock, and the ONU port adopts the SAToP recovery clock.

Figure 32-14 Line clock synchronization



Hardware configuration of the system:

- Board configuration: GPBD, SCU, TOPA (configured with the EH1A daughter board) or TOPA (configured with the CSSA daughter board).
- Clock configuration: The stratum-3 clock unit is an optional configuration. When the stratum-3 clock unit is configured, the clock quality does not meet the G.813 requirements; the quality of the clock recovered from the SAToP auto-adaptation algorithm meets only the G.8261 CES/G.823 traffic requirements. When the stratum-3 clock unit is not configured, the clock quality can meet the G.8261 CES/G.823 traffic requirements.

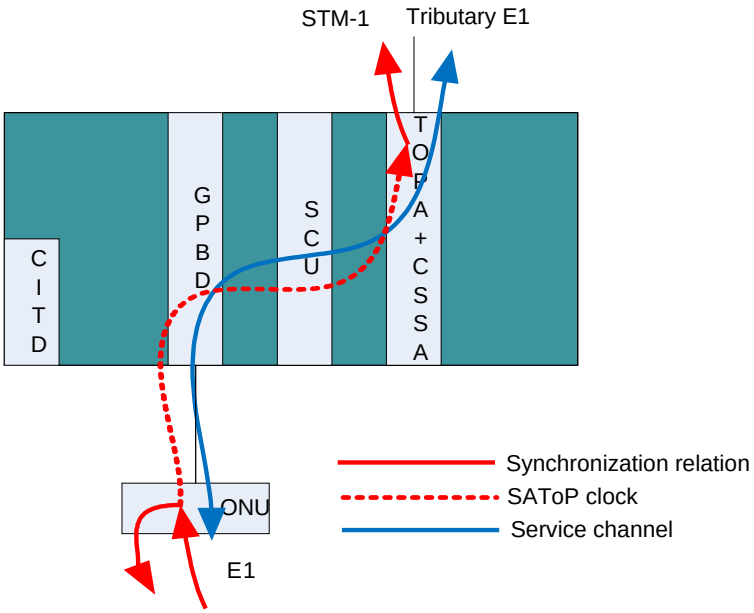
Key points of the synchronization configuration:

- Configure the system clock source of the MA5600T/MA5603T/MA5608T. For example, select the BITS clock, the STM-1 line clock, or the oscillator as the system clock.
- Select the E1 line clock as the E1 transmit clock of the TOPA board. For the STM-1 port, select the tributary E1 transmit clock as the tributary line clock. Therefore, the device interconnected with the E11 port of the TOPA board is required to serve as the master device.
- The transmit clock of the GPON port of the OLT is the system clock by default. The ONU synchronizes with the OLT through the PON port recovered line clock. In the SAToP clock recovery mode, the OLT and the ONU need not always be synchronous.
- The E1 transmit clock of the ONU adopts the SAToP recovered clock.
- The device connected to the E1 port of the ONU serves as the slave device, locking the E1 transmit clock of the ONU.

Clock mode 3: recovered clock

The recovered clock refers to the E1 clock of the ONU recovered by the OLT in the SAToP mode. In actual application, it is rare to recover the clock of a lower-level device. It applies only to certain special scenarios. The following figure illustrates the GPON device recovering the clock of the ONU in the SAToP mode. The TOPA board of the MA5600T/MA5603T/MA5608T supports one SAToP recovered clock. The TOPA board can recover the E1 clock of the ONU from the upstream SAToP data. The recovered clock can serve as the E1 transmit clock of the TOPA board for implementing synchronization between the upstream E1 port of the TOPA board and the remote E1 port of the OLT, as shown in Figure 32-15.

Figure 32-15 Recovered clock



Hardware configuration of the system:

- Board configuration: GPBD, SCU, TOPA (configured with the EH1A daughter board) or TOPA (configured with the CSSA daughter board).
- Clock configuration: The stratum-3 clock unit is an optional configuration. When the stratum-3 clock unit is configured, the clock quality does not meet the G.813 requirements; the quality of the clock recovered from the SAToP auto-adaptation algorithm meets only the G.8261 CES/G.823 traffic requirements. When the stratum-3 clock unit is not configured, the clock quality can meet the G.8261 CES/G.823 traffic requirements.

Key points of the synchronization configuration:

- Configure the system clock source of the MA5600T/MA5603T/MA5608T. For example, select the BITS clock, the STM-1 line clock, the synchronization Ethernet line clock, or the oscillator as the system clock.
- The TOPA board of the MA5600T/MA5603T/MA5608T supports only one SAToP recovered clock. Hence, select the SAToP recovered clock of an E1 port as the SAToP recovered clock source.

- Select the SAToP recovered clock source as the E1 transmit clock of the TOPA board. For the STM-1 port, select the tributary E1 transmit clock as the SAToP recovered clock source. Therefore, the device interconnected with the E1 port of the TOPA board is required to serve as the slave device, locking the E1 transmit clock of the TOPA board.
- The transmit clock of the GPON port of the OLT is the system clock by default. The ONU synchronizes with the OLT through the PON port recovered line clock. In the SAToP clock recovery mode, the OLT and the ONU need not always be synchronous.
- The ONU locks the E1 transmit clock of the lower-level device, and the ONU adopts the line clock as its E1 transmit clock.

Clock Synchronization of the Native TDM Service/SAToP Service Conversion

The MA5600T/MA5603T/MA5608T supports the conversion between the native TDM service and the SAToP service. The native TDM service is carried through GPON, and the SAToP service is carried through PW. The native TDM service/SAToP service conversion is mainly implemented by the CSPA board, which supports the conversion of 64 E1s.

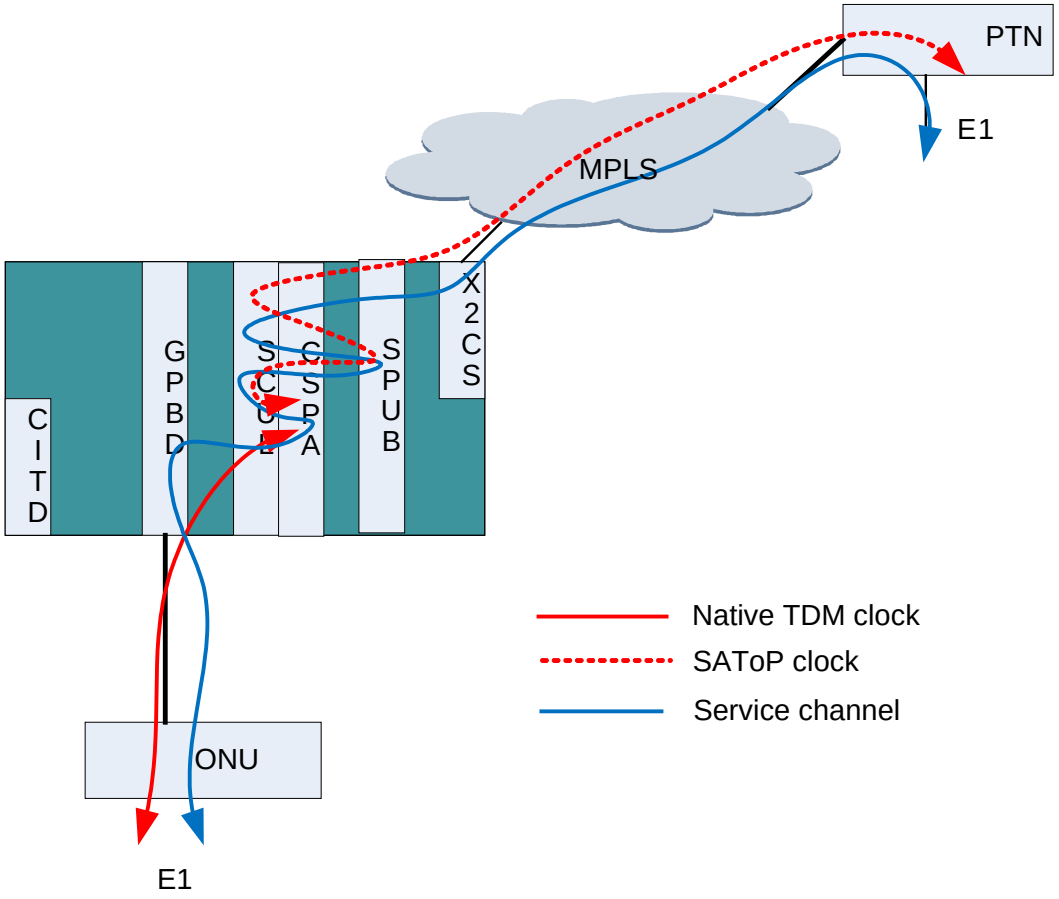
In the upstream direction, the ONU encapsulates the E1 data into VC12 and GEM frames and forwards the frames to the CSPA board of the OLT. The CSPA board decapsulates the TDMoGEM frame into E1 data, and at the same time can recover the E1 upstream clock of the ONU. Then, the CSPA board encapsulates the E1 data into SAToP data, which is then encapsulated by the SPUB board into MPLS packets and then forwarded to the PTN device.

In the downstream direction, the CSPA board receives the SAToP data from the PTN device, decapsulates the data in the E1 format, and then encapsulates the E1 data into VC12 and GEM frames, which are sent to the ONU through GPON. Then, the ONU decapsulates the VC12 and GEM frames into E1 data. The CSPA board can select one E1 from the 64 E1s for receiving the SAToP data and recover the downstream clock of the E1.

It is suggested that the clock mode be set when the CSPA board performs the native TDM service/SAToP service conversion. This mainly refers to the mode of the transmit clock in the SAToP to native TDM direction. The clock mode can be the master mode or the slave mode, and the master mode can be the system clock mode or the recovered clock mode.

The clock in the native TDM to SAToP direction always adopts the E1 clock that is recovered when the CSPA board decapsulates the GEM frames. Figure 32-16 shows the conversion between native TDM and SAToP.

Figure 32-16 Conversion between native TDM and SAToP



Master Mode

When the master mode is adopted, the CSPA board can select the system clock or the SAToP clock as the transmit clock of the native TDM service.

System clock: In this mode, the system clock of the MA5600T/MA5603T/MA5608T is used as the transmit clock of the native TDM service. The devices interconnected with the downstream E1 (ONU E1) are required to serve as the slave devices and use the line receive clock as the transmit clock, thus realizing clock synchronization of the entire system.

SAToP clock: The CSPA board can select one E1 from the 64 E1s for receiving the SAToP data and recovering the clock, which is the downstream clock of an E1 of the PTN device. The CSPA board uses the recovered clock as the transmit clock on the native TDM side. Therefore, the device interconnected with the PTN E1 is required to serve as the master device and send the transmit clock to downstream; the device interconnected with the ONU E1 serves as the slave device and uses the line receive clock as the transmit clock. In this way, end-to-end synchronization is implemented.

Hardware configuration of the system:

- Board configuration: SCU, CSPA, SPUB, and GIU (such as X2CS, GICK, GSCA, and GICD).
- Clock configuration: The stratum-3 clock unit is an optional configuration. When the stratum-3 clock unit is configured, the clock quality does not meet the G.813

requirements; the quality of the clock recovered from the SAToP auto-adaptation algorithm meets only the G.8261 CES/G.823 traffic requirements. When the stratum-3 clock unit is not configured, the clock quality can meet the G.8261 CES/G.823 traffic requirements.

Key points of the synchronization configuration:

- Configure the system clock source of the MA5600T/MA5603T/MA5608T. For example, select the BITS clock, the STM-1 line clock, the synchronization Ethernet line clock, or the oscillator as the system clock.
- Configure the clock of the CSPA board to the master mode, and select the system clock or the SAToP clock as the transmit clock to the native TDM direction.
- If the system clock is selected, the device interconnected with the ONU E1 and the PTN E1 are required to serve as the slave devices and use the line receive clock as the transmit clock, thus realizing clock synchronization of the entire system. If the SAToP clock is selected, the device interconnected with the PTN E1 is required to serve as the master device and send the transmit clock to downstream; the device interconnected with the ONU E1 serves as the slave device and uses the line receive clock as the transmit clock. In this way, end-to-end synchronization is implemented.
- By default, the system clock is used as the line transmit clock of the GPBD board, the ONU line clock synchronizes with the clock of the MA5600T/MA5603T/MA5608T, and the transmit clock of the ONU E1 port adopts the bit stream recovered clock.
- The transmit clock in the CSPA board to SAToP direction need not be configured. By default, it adopts the clock recovered from the native TDM.

Slave Mode

In the slave mode, the CSPA board adopts the clock recovered from the native TDM as the clock source, which serves as the transmit clock on the native TDM side and the SAToP side. Hence, the device interconnected with the PTN E1 is required to serve as the slave device and adopt the recovered clock for realizing the clock synchronization of the entire network.

Hardware configuration of the system:

- Board configuration: SCU, CSPA, SPUB, and GIU (such as X2CS, GICK, GSCA, and GICD).
- Clock configuration: The stratum-3 clock unit is an optional configuration. When the stratum-3 clock unit is configured, the clock quality does not meet the G.813 requirements; the quality of the clock recovered from the SAToP auto-adaptation algorithm meets only the G.8261 CES/G.823 traffic requirements. When the stratum-3 clock unit is not configured, the clock quality can meet the G.8261 CES/G.823 traffic requirements.

Key points of the synchronization configuration:

- Configure the system clock source of the MA5600T/MA5603T/MA5608T. For example, select the BITS clock, the STM-1 line clock, the synchronization Ethernet line clock, or the oscillator as the system clock.
- Configure the clock of the CSPA board to the slave mode.
- By default, the system clock is used as the line transmit clock of the GPBD board, the ONU line clock synchronizes with the clock of the MA5600T/MA5603T/MA5608T, and the transmit clock of the ONU E1 port adopts the bit stream recovered clock.
- Adopt the line recovered clock on the native TDM side as the transmit clock from the CSPA board to the native TDM. The device interconnected with the ONU E1 is required

to serve as the master device and provide the clock for the upstream direction; the device interconnected with the PTN E1 serves as the slave device and adopts the recovered clock.

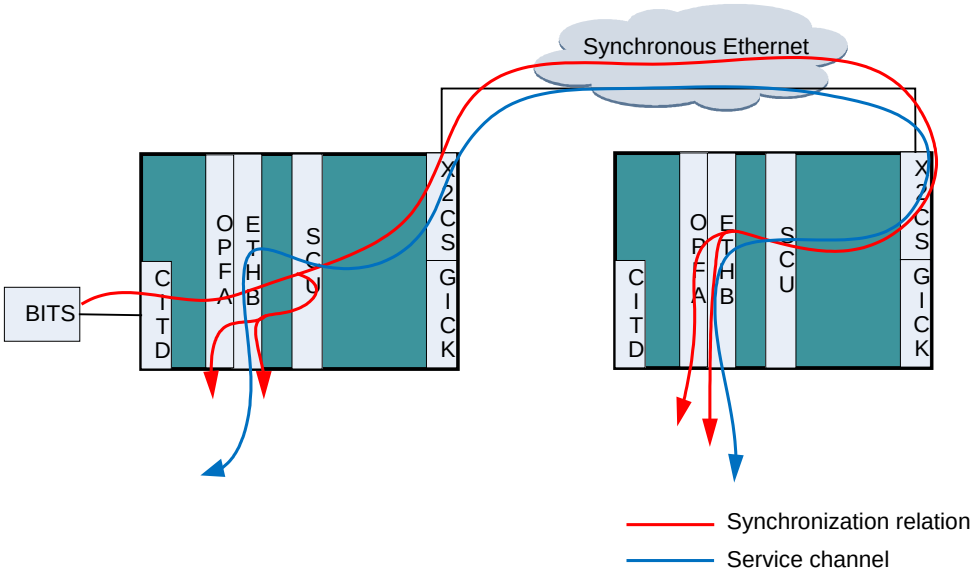
- The transmit clock in the CSPA board to SAToP direction need not be configured. By default, it adopts the clock recovered from the native TDM.

Clock Synchronization of the Synchronization Ethernet Service

Traditional Ethernet application does not consider the synchronization requirement. The Ethernet ports adopt the ±100 ppm local oscillator as the transmit clock, and the transmit clocks of the NEs are independent of each other. As such, the clocks are not precise enough. Synchronization Ethernet is a technology that recovers the clock from the bit streams on the Ethernet link and implements synchronization between Ethernets. The implementation mode is similar to the synchronization mode of the SDH/PDH networks. In the transmit direction, the high-precision system clock is adopted as the transmit clock, which is recovered and obtained at the receive end. The transmission and reception are performed by the physical layer independently, which in terms of function is compatible with traditional Ethernet.

The MA5600T/MA5603T/MA5608T supports the 10GE and GE synchronization Ethernet application, and can issue the GE and FE system clock signals. In the transmit direction, the system clock is adopted as the port transmit clock by default, the clock mode of which cannot be changed. In the receive direction, each port recovers the line clock, which serves as an optional clock source of the system.

Figure 32-17 Clock synchronization of the synchronous Ethernet service



Key points of the synchronization configuration:

- Configure the system clock source of the MA5600T/MA5603T/MA5608T. For example, select the BITS clock, the STM-1 line clock, the synchronization Ethernet line clock, or the oscillator as the system clock.
- By default, the transmit clock of the upstream port and access port of the MA5600T/MA5603T/MA5608T is adopted as the system clock source. The clock mode cannot be changed and does not need any extra configuration.

- The line recovered clock can serve as the system clock source. The line recovered clock should be configured in such a way that mutual-locking between NEs is prevented.
- When lower-level devices need to synchronize with the network clock, the lower-level devices should lock the line clock of the GPBD boards.

32.5.3 Configuring the Physical Clock

On a digital network comprising the MA5600T and other devices, the primary problem to be solved is clock synchronization for carrying the traditional TDM service. To ensure the system clock synchronization of each device in the digital network, a system clock source must be specified.

Context

IP-based solution is the trend of future network and service development, so is the trend of the bearer network. Difficulties, however, currently exist in the transition from the SDH-based traditional network to the IP-based Ethernet bearer network. One key technology involved is how to carry traditional TDM service on the new network. Traditional TDM service has two major applications: voice service and clock synchronization service. In a traditional communications network architecture, the TDM service of the fixed network is mainly voice service. Cumulative inconsistency between the clocks at both ends of the bearer network over a long time causes frame slip. On a communications network, the wireless application has the most rigorous requirements on the clock frequency. The frequencies of different base stations must be synchronized within a specified precision. Otherwise, re-sync occurs during the base station switching.

To ensure clock synchronization among devices, relevant clock synchronization methods are adopted based on the clock source solution provided by an upper-layer device.

Table 32-5 Clock configuration method

Configuration Method	Configuration Principle
Configuring the system clock based on the priority	This configuration method is similar to the basic configuration of the clock. If a device has multiple clock sources and the precision of the clock sources is provided, the clock source with the highest precision is generally configured with the highest priority.
Configuring the system clock based on the SSM clock source selection mode	This configuration method is adopted if the clock transmitted from an upper-layer device contains SSM information and all clock sources are selected based on the SSM clock source selection mode.
Configuring external clock	This configuration method is adopted to output an independent clock signal from the CITD BITS OUT interface to serve as the clock source of a lower-layer device.

Configuring the System Clock Based on the Priority

If a device has multiple clock sources and the precision of the clock sources is provided, you need to configure the priorities of the clock sources. Generally, the higher the precision is, the higher the priority is.

Context

A clock source can be an external BITS clock or a line clock from the upper-layer node. The clock module automatically judges the types of the specified clock sources (BITS, TDM, or SDH), and sends them according to their priorities to the clock module, serving as clock sources for phase lock.

NOTE

When the SSM signal is used as the input clock signal and the system selects the clock source based on the SSM signal, see Configuring the System Clock Based on the SSM Clock Source Selection Mode.

Procedure

Run the **clock source sourceid { frameid/slotid/portid [bits-clktype bits-impedance]** command to configure the system clock source.

Specify the clock signals extracted from a certain port as the system clock source.

- The system supports 10 clock sources in total.
- Only the external clock sources on the physical entities are added by running this command and the external clock sources are numbered. To enable the relevant external clock source, you need to run the **clock priority** command to determine whether the relevant clock source is available.
- The system clock cannot serve as the system clock source.

Step 1 Run the **clock priority system p0/p1/p2/p3/p4/p5/p6/p7/p8/p9** command to configure the priority of the system clock source.

- The system supports 10 clock source priorities. The highest priority is p0 and the lowest priority is p9.
- When the clock source is selected based on the priority, the system does not check the quality of the clock source. Therefore, you must configure the clock source of high quality with a high priority.
- After the priority of the clock source is configured, the system selects the clock source with the highest priority and in the normal state as the system clock source.
- When the clock source with the highest priority is faulty, the system automatically switches to the clock source with the second highest priority.
- When the clock source with the highest priority recovers, the system switches back to this clock source.

----End

Example

Assume the following configurations: On the MA5600T, obtain three clock sources from port BITS on the CITD board and ports 0/5/0 and 0/5/1 of the TOPA board as the clock source 0, clock source 1, and clock source 2 of the system. Configure clock source 0 with the highest

priority and configure clock source 2 with the lowest priority. To perform the preceding configurations, do as follows:

```
huawei(config)#clock source 0 0/0/0 2MHz 120ohm
huawei(config)#clock source 1 0/5/0
huawei(config)#clock source 2 0/5/1
huawei(config)#clock priority system 0/1/2
```

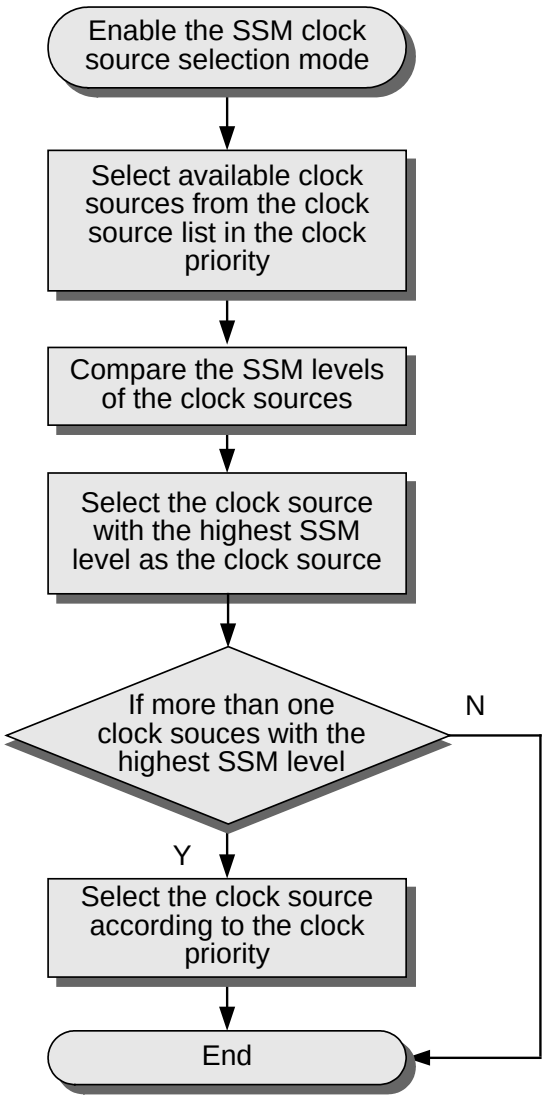
Configuring the System Clock Based on the SSM Clock Source Selection Mode

If the clock transmitted from an upper-layer device contains a synchronization status message (SSM) and all clock sources are selected based on the SSM, you need to configure the system clock based on the SSM clock source selection mode.

Context

By default, the SSM clock source selection mode is disabled. That is, the system selects the clock source based on the priority. The system enables the SSM clock source selection mode only after the system determines that the clock source contains an SSM and the entire system are based on the SSM clock source selection mode.

For the detailed SSM clock source selection process, see the following flowchart.



Procedure

Enable the SSM clock source selection mode.

Run the **clock ql-mode enable** command to enable the SSM clock source selection mode.

Step 1 Run the **clock source sourceid { frameid/slotid/portid [bits-clktype bits-impedance]** command to configure the system clock source.

Step 2 Run the **clock priority system p0/p1/p2/p3/p4/p5/p6/p7/p8/p9** command to configure the clock source range based on the SSM clock source selection mode and the priority sequence of the corresponding clock sources if the SSM quality levels are the same.

Step 3 (Optional) Run the **clock ql sourceid clock-ql** command to configure the SSM quality level for the clock source. If the SSM clock source selection mode is enabled, but a clock source does not support the output of an SSM, you need to manually configure the SSM quality level for the clock source. After the SSM quality level is configured, the device no longer matches the received SSM.

NOTE

- When the SSM clock source selection mode of the system is disabled, the SSM quality level of the clock source cannot be set.
- When the system selects the clock source based on the SSM quality level, the system selects the clock sources based on the priority and then compares the SSM of the clock sources. Finally, the system selects the clock source with the highest SSM quality level as the system clock source. If there are multiple clock sources with the same SSM qualify level, the system selects the clock source based on the priority.

Step 4 (Optional) Run the **clock ql input lower-limit** command to configure the lowest synchronization status message (SSM) quality level threshold of the clock source. When the SSM quality level of the clock source of the upper-layer device is higer than or equal to that of the device, the clock of the upper-layer device is traced. Otherwise, the clock of the device is in the holdover state and switched to the free-run state after 24 hours elapses.

Step 5 (Optional) Run the **clock ql output** command to configure whether the specified port sends the SSM quality level.

----End

Example

Example 1:

Assume the following configurations: Configure the SSM clock source selection mode as the system clock source selection mode. Obtain three clock sources from ports 0/20/0 and 0/20/1 of the X2CS board and port 0/19/0 of the GICK board as the clock source 0, clock source 1, and clock source 2 with the SSM. Configure clock source 0 with the highest priority, clock source 2 with the lowest priority, and the lowest threshold for the received clock sources are QL-SSU-B. To perform the preceding configurations, do as follows:

```
huawei(config)#clock ql-mode enable
huawei(config)#clock source 0 0/20/0
huawei(config)#clock source 1 0/20/1
huawei(config)#clock source 2 0/19/0
huawei(config)#clock priority system 0/1/2
huawei(config)#clock ql input threshold  lower-limit QL-SSU-B
```

Example 2:

Assume the following configurations: Configure the SSM clock source selection mode as the system clock source selection mode. Obtain clock sources with SSM from ports 0/19/0 and 0/19/1 of the GICK board as clock source 0 and clock source 1; and obtain a clock source that does not support SSM from port 0/0/0 of the CITD board as clock source 2. Configure clock source 0 with the highest priority and configure clock source 2 with the lowest priority.

Set the SSM of clock source 2 to QL-SSU-A and the lowest threshold for the received clock sources is QL-SSU-B.

Send SSM from port 0/3/1. To perform the preceding configurations, do as follows:

```
huawei(config)#clock ql-mode enable
huawei(config)#clock source 0 0/19/0
huawei(config)#clock source 1 0/19/1
huawei(config)#clock source 2 0/0/0 2MHz 120ohm
huawei(config)#clock ql 2 QL-SSU-A
huawei(config)#clock priority system 0/1/2
```

```
huawei(config)#clock ql input threshold  lower-limit QL-SSU-B
huawei(config)#clock ql output 0/3/1 enable
```

Configuring External Clock

The MA5600T can select a system clock output or export the line clock as the clock source of another device.

Context

The external clock output of the MA5600T supports the selection of the following two benchmark clocks.

- Select the system clock as the output benchmark clock.
- Select the line clock as the output benchmark clock.

Procedure

Run the **clock external mode** command to configure the switch mode of the output clock.

In the auto-trace mode, enable the output clock or disable the output clock depending on the system.

- In the case of the SSM clock source selection mode, if the SSM quality level of the input clock is not lower than the threshold, the output clock port (T4) is enabled; otherwise, the output clock port is disabled.
- In the case of the priority clock source selection mode, the output clock is enabled if there is a clock source serving as the current source of the output clock. Otherwise, the output clock is disabled.

In the fix-trace mode, the system clock is output fixedly.

In the no-trace mode, the output clock is manually disabled.

NOTE
The system defaults to the fix-trace mode.

Step 1 Run the **clock external bits-type { 2MHz |2Mbps }** command to set the signal type of the BITS port on the CITD board to 2 MHz or 2 Mbit/s.

NOTE
The system defaults to 2 Mbit/s.

Step 2 Configuring the clock source of the external clock

1. Run the **clock source sourceid { system | { frameid/slotid/portid [1588 | bits-clktype bits-impedance]** command to configure the clock sources of the external clock.

NOTE
The CITD port does not support clock sources of the external clock.

2. Run the **clock priority external p0/p1/p2/p3/p4/p5/p6/p7/p8/p9** command to configure the clock sources and their priorities.

Step 3 (Optional) Run the **clock external output threshold clock-ql** command to configure the threshold of the SSM quality level for the output clock.

- If the output clock is in the auto-trace mode and its SSM quality level is lower than the threshold, the output clock is disabled automatically.

- The threshold takes effect only when the clock source mode is the SSM clock source selection mode. To configure the threshold, see Configuring the System Clock Based on the SSM Clock Source Selection Mode.

----End

Example

To obtain synchronization Ethernet clock sources from port 0/19/0 of the GICK board and port 0/20/1 of the X2CA board as system clock source 1 and system clock source 2, configure their priorities, and set the output clock to auto-trace, do as follows:

```
huawei(config)#clock external mode auto-trace
huawei(config)#clock source 1 0/19/0
huawei(config)#clock source 2 0/20/1
huawei(config)#clock priority external 2/1
```

32.5.4 Physical Layer Clock/Time Synchronization Standards and Protocols Compliance

The following lists the reference standards and protocols of this feature:

- ITU-T G.813 Timing characteristics of SDH equipment slave clocks (SEC)

  ITU-T G.813 defines the requirements on synchronization by SDH devices. The native TDM, CESoP, and ATM services provided by the MA5600T/MA5603T/MA5608T are carried through the STM-1 port. Therefore, the ITU-T G.813 requirements on the clock feature must be met. ITU-T G.813 specifications include the following items. To meet the ITU-T G.813 specifications, the MA5600T/MA5603T/MA5608T must be configured with the CKMC stratum-3 clock daughter board.
  - Items defined by the G.813 Recommendations:
    - Frequency accuracy
    - Pull-in and pull-out ranges
    - Wander generation
    - Jitter output
    - Input wander tolerance
    - Input jitter tolerance
    - Noise transfer
    - Short-term phase transient response
    - Long-term phase transient response (holdover)
    - Phase response to input signal interruptions
    - Phase discontinuity
- ITU-T G.823 (the control of jitter and wander within digital networks which are based on the 2048 kbit/s hierarchy)

  The ITU-T G.823 describes the jitter and wander of the PDH interface in the 2048 kbit/s synchronous system. Here, the E1port is mainly referred to, which meets the G.824 requirements for the 1544 kbit/s system. G.823 defines the requirements on the traffic interface and the synchronization interface. The traffic interface performance is the basic requirement for service transmission. To transmit synchronous clock signals, the E1 port

must meet the requirements of the synchronization interface. The synchronization interface has stricter requirements on jitter and wander than the traffic interface does.

  - The G.823 defines the following items for the traffic interface and the synchronization interface:
    - Jitter output
    - Output wander
    - Input jitter and wander tolerance

- ITU-T G.736, Characteristics of a synchronous digital multiplex equipment operating at 2048 kbit/s

  The ITU-T G.736 defines the jitter transfer characteristics of the E1 port.

- ITU-T G.825, The control of jitter and wander within digital networks which are based on the synchronous digital hierarchy (SDH)

  The ITU-T G.825 defines the jitter and wander characteristics of SDH devices.

  - The ITU-T G.825 defines the following items for the STM port:
    - Jitter output
    - Output wander
    - Input wander tolerance
    - Input jitter tolerance
    - Jitter and wander generation
    - Jitter and wander transfer

- ITU-T G.8261, Timing and Synchronization Aspects in Packet Networks

  The ITU-T G.8261 defines the wander budget of CES and synchronous Ethernet for packet networks. The ITU-T G.8261 requirements are similar to the requirements of the ITU-T G.823 on the TDM network. The CESoP service and synchronous Ethernet clock feature of the MA5600T/MA5603T/MA5608T need to meet the ITU-T G.8261 requirements.

- ITU-T G.8262, Timing characteristics of synchronous Ethernet equipment slave clock (EEC)

  The ITU-T G.8262 defines the requirements on the synchronous Ethernet clock system, which are equal to the G.813 and G.812 requirements on the TDM network. To meet the ITU-T G.8262 specifications, the MA5600T/MA5603T/MA5608T must be configured with the CKMC stratum-3 clock daughter board.

- ITU-T G.703, Physical/Electrical characteristics of hierarchical digital interfaces

  The E1 and BITS ports support the 75-ohm or 120-ohm impedance setting, and output physical signal template and rate that meet the ITU-T G.703 Recommendations.

- ITU-T G.8264, Distribution of timing through packet networks

  The ITU-T Recommendation G.8264 defines the SSM protocol and message format for the synchronous Ethernet to ensure clock synchronization between Ethernet devices.

32.6 1588v2

The IEEE organization proposed the IEEE 1588v2 precision time synchronization protocol, which supports system-wide synchronization accuracy within the sub-microsecond range. In contrast to the GPS solution, the IEEE 1588v2 solution achieves the same time precision but has advantages in terms of cost, maintenance, and security. It has become the most popular time synchronization protocol in the industry.

32.6.1 Why Is 1588v2 Required

Background

Traditionally, time synchronization chains use the NTP, which can only ensure time precision within milliseconds. This precision is far below the time synchronization precision of microseconds required by wireless base stations.

Therefore, GPS is currently used at wireless base stations for frequency and time synchronization. When the GPS is used for time synchronization, there are the following problems:

- Difficult to select a site and install devices.
- Difficult to maintain. The fault rate of a GPS is high (over 1%).
- Difficult to route feeder cables.
- Highly risky.
- High costly.

To accommodate the high precision requirement of wireless base stations and resolve the problems of the GPS solution, operators are eager for a comprehensive and high-precision time synchronization solution.

IEEE 1588v2 high-precision Time Transfer Protocol

Based on the synchronization requirements for different wireless systems and time synchronization condition, IEEE defines IEEE 1588v2 high-precision time protocol (full name: Precision Clock Synchronization Protocol for Networked Measurement and Control Systems and abbreviation: PTP). The protocol supports time synchronization precision in the sub-microsecond range.

In contrast to the GPS solution, IEEE 1588v2 solution achieves the same time precision but has advantages in terms of cost, maintenance, and security. It has become the most popular time transfer protocol in the industry because of the following advantages:

- Space localization: The IEEE 1588v2 protocol is applicable to local area networks supporting multicast messaging including but not limited to Ethernet.
- Low cost: The IEEE 1588v2 protocol helps minimize the demand for network resources and CPU resources and therefore enables low-cost applications.
- Network transformation trend: IEEE 1588v2 is suitable for future integrated networks over IP.

32.6.2 1588v2 Basic Concepts

Device Model

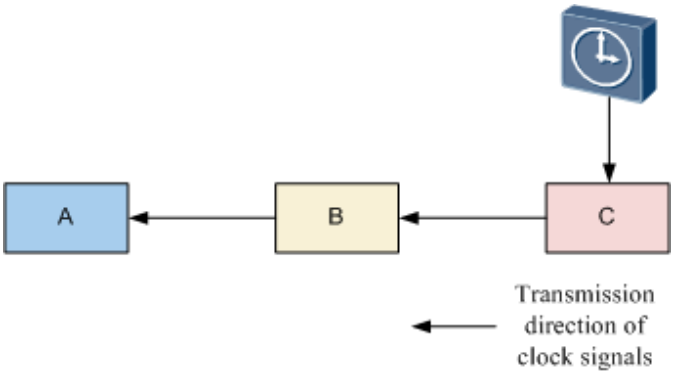
Device Model

When a pair of nodes perform time synchronization, the upstream node distributing the reference time signals is the master node and the downstream node receiving the reference time signals is the slave node. As shown in Figure 32-18, A is the master node, and B is the slave node.

In the subnet system, grand master is the master clock. Each system has only one grandmaster clock (GMC), and each subnet has only one master clock. Slave clocks are kept synchronized with the master clock.

As shown in Figure 32-18, C is the grand master, and nodes A and B must be synchronized with C.

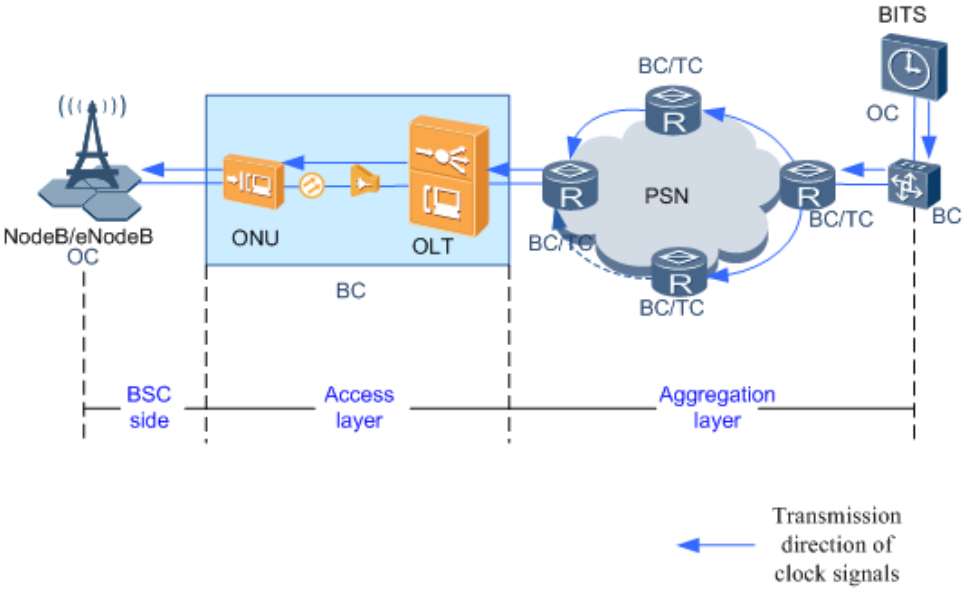
Figure 32-18 Synchronization on devices



1588v2 Device Model

The IEEE 1588v2 standard defines five network node models: ordinary clock (OC), boundary clock (BC), end-to-end transparent clock (E2E TC), peer-to-peer transparent clock (P2P TC), and management node.

Figure 32-19 Device model



OC

An OC has only one 1588v2 clock interface (a clock interface enabled with 1588v2) through which the OC synchronizes with an upstream node or distributes time signals to downstream

nodes. An OC has only one physical port. Therefore, in practice, an OC can either be a grandmaster (GM) clock or a slave lock.

As shown in Figure 32-19, BITS is generally configured as an OC, functioning as the GM clock in the entire network; a basic station, as a slave device, is also configured as an OC.

BC

A BC has several physical ports for network communication. Each port is similar to the port of an OC, which is used to connect multiple domains. One of the port is used to synchronize with an upstream node. The other interfaces can be used to distribute time signals to downstream nodes.

In practice, a BC synchronizes its time with the upstream device and issues the synchronized time to the downstream devices.

As shown in Figure 32-19, BC contains the devices (access network devices OMU and OLT) at the aggregation layer.

E2E TC

An E2E TC provides several physical ports, forwards all PTP messages, and measures and corrects the residence time of PTP event messages traversing the E2E TC.

P2P TC

A P2P TC has multiple ports. In addition to functions of an E2E TC, a P2P TC is also able to compute and correct the link delay between each port and similarly equipped port on another node sharing the link, that is, the link peer.

Difference between a TC and BC/OC: BC/OC must be synchronized with the clock on other devices, and TC does not restore the clock from IEEE 1588 packets but processes and forwards IEEE 1588 packets.

As shown in Figure 32-19, a TC can function as a device at the aggregation layer.

Management node

A management node has multiple ports and serves as a management interface to PTP management messages. The management node is only used to manage the synchronization nodes and does not provide the synchronization function.

Clock Domain and Clock ID

Clock Domain

Logically, a physical network can be divided into multiple clock domains. Each clock domain has a reference time with which all devices in the domain are synchronized. Each clock domain has its own reference time and these times are independent of one another.

A device can join only one clock domain and can synchronize only with the synchronization time of that clock domain.

Clock ID

A clock ID identifies a clock in an IEEE 1588v2 clock subnet. In an IEEE 1588v2 packet, a clock source ID occupies eight bytes. It consists of two parts: The first three bytes are OUI code, and the other five bytes are extended ID.

Organizational Unique Identifier (OUI): an organization identifier uniformly assigned by the IEEE standard.

Extended ID: an identifier uniformly assigned by the organization represented by the OUI to ensure that the clock ID in each IEEE 1588v2 packet is unique.

1588v2 Messages

Message Type

As show in Table 32-6, there are two types of IEEE 1588v2 messages: event message and general message.

Table 32-6 Message Type

Message Type	Message	Description
Event message	• Sync • Delay_Req • Pdelay_Req • Pdelay_Resp	Event messages are timed messages in that an accurate timestamp is generated both at the device ingress and egress. Timestamps need to be processed in a timely manner at the time of message transmission and reception.
General Message	• Announce • Follow_Up • Delay_Resp • Pdelay_Resp_Follow_Up • Management • Signaling	General messages are not timed messages and do not require accurate timestamps. Timestamps do not need to be processed in a timely manner at the time of message transmission and reception.

- The Sync, Delay_Req, Follow_Up, and Delay_Resp messages are used to generate and communicate the timing information needed to synchronize OCs and BCs using the delay request-response mechanism.
- The Pdelay_Req, Pdelay_Resp, and Pdelay_Resp_Follow_Up messages are used to measure the link delay between two clock ports implementing the Pdelay mechanism. The link delay is used to correct timing information in Sync and Follow_Up messages in systems composed of P2P TCs. OCs and BCs that implement the Pdelay mechanism can synchronize using the measured link delays and the information in the Sync and Follow_Up messages.
- The Announce message is used to establish the synchronization hierarchy. That is, information related to the best master clock (BMC) algorithm is carried in the Announce messages.
- The management messages are used to query and update the PTP data sets maintained by clocks. These messages are also used to customize a PTP system and for initialization and fault management. Management messages are used between management nodes and clocks.

- The signaling messages are used for communication between clocks for all other purposes. For example, signaling messages can be used for negotiation of the rate of unicast messages between a master clock and its slave clocks.

Message Format

An IEEE 1588v2 message must have a header, body, and suffix. The suffix length may be 0. An IEEE 1588v2 message may be encapsulated in Ethernet or IP format.

IEEE 1588v2 message encapsulated in Ethernet format

Messages encapsulated in Ethernet format are classified into IEEE 1588v2 over Ethernet messages without VLAN tags and IEEE 1588v2 over Ethernet messages with VLAN tags.

Figure 32-20 Format of an IEEE 1588v2 over Ethernet message

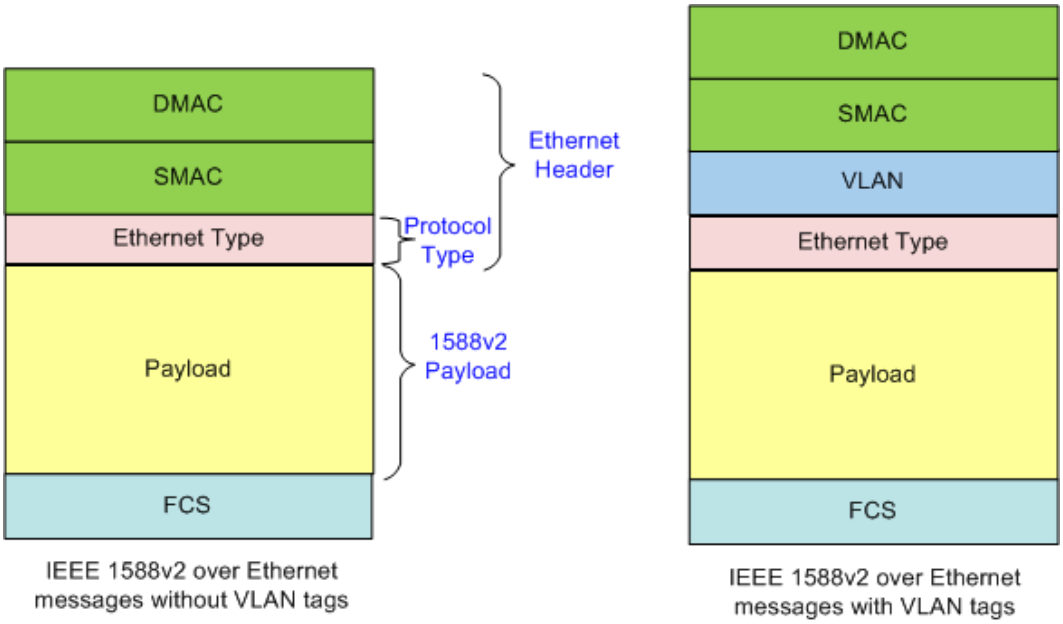


Table 32-7 describes the fields in an IEEE 1588v2 over Ethernet message.

Table 32-7 Fields in an IEEE 1588v2 over Ethernet message

Field Name	Bit Width (Unit: Bit)	Description
DMAC	6 x 8 = 48	Destination MAC address. It is configured by a user. • When the destination MAC address is not configured, the multicast encapsulation mode is used. This field is filled with 01-80-C2-00-00-0E for Pdelay packets and filled with 01-1B-19-00-00-00 for other packets. • When the destination MAC address is configured, the unicast encapsulation mode is used. This field is filled with the configured destination MAC address or with the SMAC of

Field Name	Bit Width (Unit: Bit)	Description
		received packets.
SMAC	6 x 8 = 48	Source MAC address. It is configured by a user. It is the MAC address of an NE's control board.
VLAN	4 x 8 = 32	VLAN tag. It is configured by a user. At most one VLAN tag is supported, and TPID can only be 0x8100.
Ethernet Type	2 x 8 = 16	Ethernet type. It is configured by the system. This field has a fixed value of 88F7, which indicates an IEEE 1588v2 over Ethernet message.
Payload	N = (A value ranging from 44 to 64) x 8	Payload of an IEEE 1588v2 over Ethernet message.
FCS	4 x 8 = 32	CRC32 check value.

IEEE 1588v2 message encapsulated in IP format

Messages in encapsulated in IP format are classified into IEEE 1588v2 over IPv4 messages without VLAN tags and IEEE 1588v2 over IPv4 messages with VLAN tags.

Figure 32-21 Format of an IEEE 1588v2 over IPv4 message

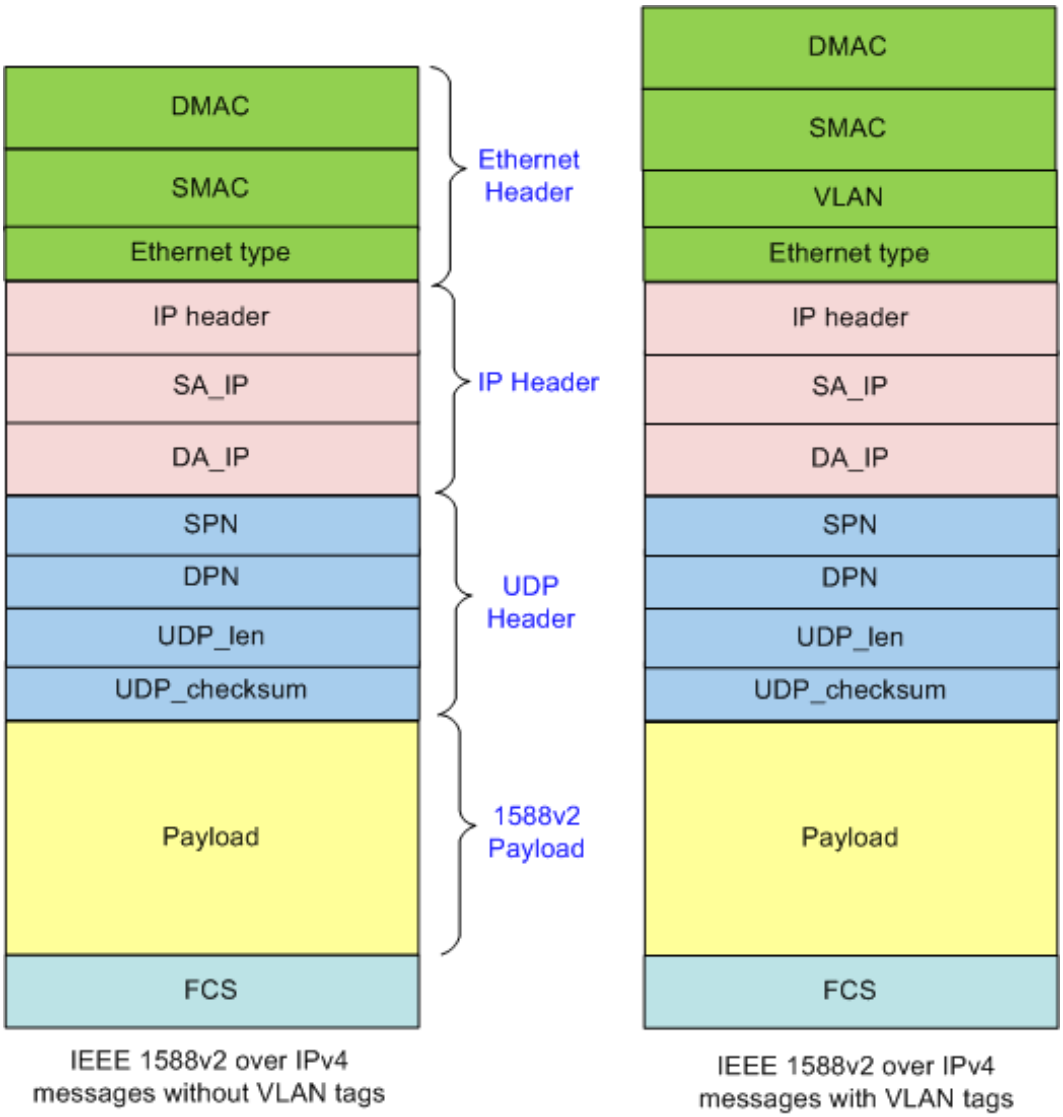


Table 32-8 lists the fields in an IEEE 1588v2 over IPv4 message.

Table 32-8 Fields in an IEEE 1588v2 over IPv4 message

Field Name	Bit Width (Unit: Bit)	Description
DMAC	6 x 8 = 48	-
SMAC	6 x 8 = 48	-
Ethernet	2 x 8	-

Field Name	Bit Width (Unit: Bit)	Description
Type	= 16	
VLAN	4 x 8 = 32	-
IP header	12 x 8 = 96	An IP header consists of: IP version (4 bits) Header length (4 bits) Service type (8 bits) Total length (16 bits) Identifier (16 bits) Flag (3 bits) Fragment offset (13 bits) Lifetime (8 bits) Protocol (8 bits) Header checksum (16 bits)
SA_IP	4 x 8 = 32	Source IP Address. If an NE port is configured with an IP address, this field is filled with the NE's IP address. If an NE port is not configured with an IP address, this field is filled with the NE's IP address.
DA_IP	4 x 8 = 32	Destination IP Address. When the destination IP address is not configured, the multicast encapsulation mode is used. This field is automatically filled with the destination IP address. It is filled with 224.0.0.107 for Pdelay packets and filled with 224.0.1.129 for other packets. When the destination IP address is configured, the unicast encapsulation mode is used. This field is filled with the configured destination IP address or with the SA_IP of received packets.
SPN	2 x 8 = 16	Source port ID. It is filled by the system with the ID of the NE port that transmits the packet.
DPN	2 x 8 = 16	Destination port ID, which is used to identify whether the packet is a 1588 packet. It is filled by the system.
UDP _Len	2 x 8 = 16	Data packet length. It is calculated by the system automatically using the following formula: UDP header length + 1588v2 payload length.
UDP_checksum	2 x 8 = 16	Header checksum. The value is 0 if the check is not required.
Payload	N = (A	-

Field Name	Bit Width (Unit: Bit)	Description
	value ranging from 44 to 64) x 8	
FCS	4 x 8 = 32	-

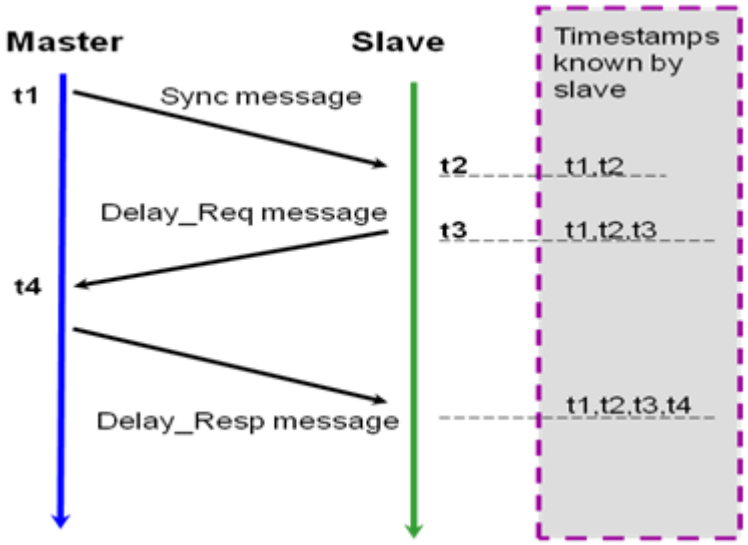
32.6.3 1588v2 Principle

Clock Synchronization Mechanism

Time Synchronization Principle

The master and slave exchanges IEEE 1588v2 messages in the following procedure:

Figure 32-22 IEEE 1588v2 time synchronization principle



1. The master sends a Sync message at t1 and carries the t1 timestamp in the Sync message.
2. The slave receives the Sync message at t2, locally generates the t2 timestamp, and extracts the t1 timestamp from the Sync message.
3. The slave sends a Delay_Req message at t3 and locally generates the t3 timestamp.
4. The master receives the Delay_Req message at t4, locally generates the t4 timestamp, and sends the Delay_Req message with the t4 timestamp back to the slave.

5. The slave extracts the t4 timestamp from the Delay_Resp message after receiving it.

"Delayms" is the path delay in the direction from the master to the slave, "Delaysm" is the path delay in the direction from the slave to the master, and "Offset" is the time offset between the slave and master. All of them are variables.

t2-t1=Delayms+Offset

t4-t3=Delaysm-Offset

When Delayms = Delaysm, that is, when the transmit and receive links between the master and slave are symmetric, the following formula is satisfied:

Offset=[(t2-t1)-(t4-t3)]/2

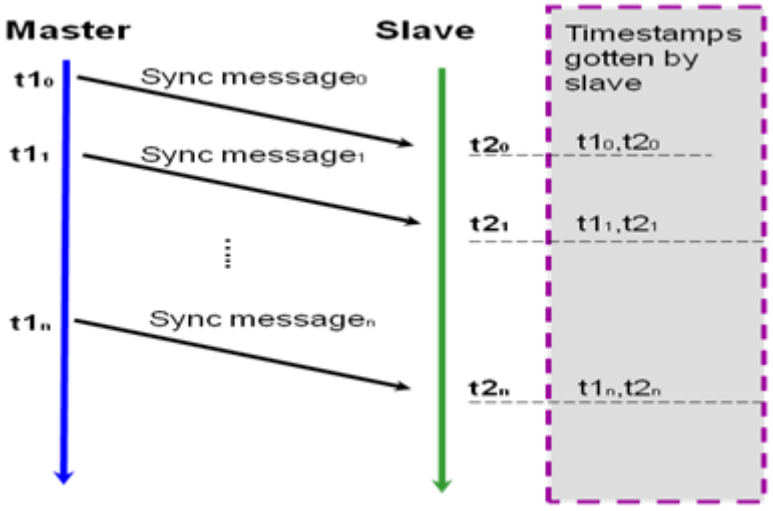
The slave can calculate the time offset between itself and the master based on the t1, t2, t3, and t4 timestamps and then corrects its own time to get synchronized with the master.

The preceding principle shows that IEEE 1588v2 time synchronization is based on the link symmetry between the master and slave. If the transmit and receive links are asymmetric, synchronization errors will occur and it will be half of the link delay asymmetry.

Frequency Synchronization Principle

The IEEE 1588v2 protocol implements frequency synchronization by exchanging Sync messages between the master and slave.

Figure 32-23 IEEE 1588v2 frequency synchronization principle



The master periodically sends Sync messages to the slave. If the slave frequency is synchronized to the master frequency, then the accumulative time errors within the same time periods are the same, as long as the path delay changes are neglected.

In other words, $t2_1-t2_0=t1_1-t1_0$, $t2_2-t2_1=t1_2-t1_1$, $t2_3-t2_2=t1_3-t1_2$,...$t2_n-t2_0=t1_n-t1_0$.

If $t2_n-t2_0$ is greater than $t1_n-t1_0$, then the slave frequency is higher than the master frequency, which means the slave frequency must be decreased. Reversely, the slave frequency must be increased.

BMC Algorithm

Background

The IEEE 1588v2 synchronization process involves two phases:

1. Establish the master/slave hierarchy.

   Within a clock domain, each port of OC and BC nodes examines the contents of all Announce messages received on the port. Using the best master clock (BMC) algorithm, the OC and BC nodes analyze the Announce message contents to determine the states of the reference sources and each port. In this manner, the OC and BC nodes establish the master/slave hierarchy and the tracing paths.

2. Frequency/time synchronization

   After the master/slave hierarchy and tracing paths are established, the master and slave exchange event messages such as Sync, Delay_Req, and Delay_Resp to implement frequency or time synchronization.

This section describes the BMC algorithm principles.

Overview

The BMC algorithm, defined in the IEEE 1588v2 protocol, is used to determine the master-slave relationship among clocks in a network. With this algorithm, clocks in the network are classified into master clock and slave clock. Slave clocks lock the frequency or time of the master clock. In the event of changes in the network or clock source attributes, the best master clock is re-selected using the BMC algorithm so that the time and clock in the entire network are synchronized.

- From the network aspect, the BMC algorithm helps form a tree structured (GM clock as the root, the best clock source) master-slave synchronization hierarchy for clocks in the entire network.
- From the node aspect, the BMC algorithm helps determine the master clock in each clock node. All Announce messages received on the ports of a given clock node and the local clock are compared to determine which clock is the best.

If the best clock is the local clock, the local clock functions as a GM clock. If the best clock is an external clock, the external clock is selected as the master clock of the local clock, that is, the local clock locks this external clock.

BMC Implementation

The BMC algorithm involves data set comparison, status decision, state machine, and data set update. The data set members include priority, clock quality (level, precision, and offset), and stepRemoved (the distance between the device and the clock source). These members are contained in ANNOUNCE packets defined by 1588v2.

An NE obtains the information about these members from the ANNOUNCE packets received on its ports. The NE uses the data set comparison algorithm to select the best external clock source (Ebest). Then the NE uses the status decision and state machine to determine the master/slave state of each port based on the Ebest, local clock D0, and data sets on the ports (Erbest).

If the NE's clock is the master clock on the entire network, the NE's all ports transmit clock signals to other NEs. If the NE's clock is the slave clock, a port on the NE must be the slave

port for tracing the master clock. The NE transmits the obtained master clock data to lower-layer NEs in ANNOUNCE packets through its master port.

Data set comparison algorithm

The algorithm compares two data sets of two ports. One data set contains the default information about the local clock, and the other data set contains the information in 1588v2 packets received from an external port. The algorithm compares the following attributes in the data sets in order to obtain the best ANNOUNCE packet:

- priority1: indicates the priority configured by a user. It identifies a clock's relative priority in the master clock set.
- clockClass: indicates the clock source's level.
- clockAccuracy: indicates the clock's precision.
- OffsetScaledLogVariance: indicates the clock's stability. It is the quality of clock signals that are received by an NE. The NE obtains the value using Allan deviation.
- priority2: indicates the priority configured by a user. It is used to select a better clock source when other attributes of clock sources are the same.
- Clock Identity: indicates the clock ID on a device.

If an NE receives, through its two different ports, two ANNOUNCE messages that are sent by the same grandmaster clock source and if the data set comparison results are the same, the distance between each port and the grandmaster clock source (that is, the number of hops between each port and the grandmaster clock source, specified by stepsRemoved in the ANNOUNCE messages) determines the port for clock tracing. The port that is closer to the grandmaster clock source (with less hops in between) is selected to trace the grandmaster clock source. This case usually occurs in a precision time protocol (PTP) system that has a loop.

Status decision algorithm

The status decision algorithm involves the data sets D0, Erbest, and Ebest. The algorithm determines the status decision code and recommended status for each port based on the data set comparison results. The recommended status is a factor for triggering the state machine. The state machine obtains a port's next status based on the triggering factor and current status. The status decision code is used to select a data set to update the global date set of the NE.

State machine

The state machine obtains a port's next status based on the triggering factor and current status. The triggering factor may be the recommended status of the port or an external event that may cause status changes. An external event can be port disabling or enabling, going online, or disconnection due to timeout.

Data set update

An NE uses the status decision code of each port to determine whether to use Ebest or D0 to update the global data set. The updated global data set is used as the tracing data set of the NE and is transmitted by the NE's master port to lower-layer NEs.

In the LTE carrying scenario, the BMC uses only the data set comparison algorithm. The algorithm is implemented as follows: Each port uses the algorithm to compare the ANNOUNCE packets that contain clock information and select the best clock source (Erbest) received on the port.

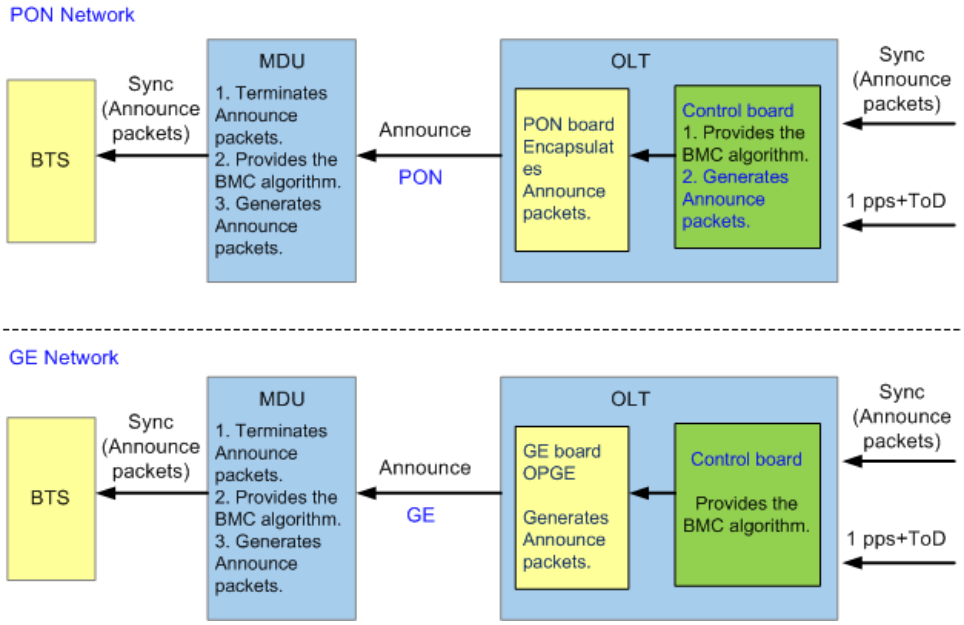
BMC Implementation in an FTTM LTE Scenario

As shown in Figure 32-24, the BMC in an FTTM LTE scenario is implemented as follows:

1. An ONT obtains ANNOUNCE packets from an uplink Ethernet port.
2. The OLT's control board obtains the clock source from the packets and compares the clock source with other obtained clock sources. Then, the control board selects the best clock source as the system time source.
3. The OLT constructs ANNOUNCE packets and broadcasts the packets using PON encapsulation mode (GEM) ports to lower-layer multi-dwelling units (MDUs).
4. An MDU obtains ANNOUNCE packets from PON lines and obtains clock source information from the ANNOUNCE packets. Then, it compares the clock source with other obtained clock sources, and selects the best clock source as the system time source.
5. An MDU constructs ANNOUNCE packets and transmits them to lower-layer base stations using user-side ports.

OLT operations in GE networking are different from those in PON networking. Specifically, in GE networking, the OLT sends the PTP packets to MDUs through GE ports after constructing an ANNOUNCE packet.

Figure 32-24 BMC implementation in an FTTM LTE scenario



Delay Compensation Mechanism

Background

Based on the principles described, IEEE 1588v2 time synchronization is based on delay symmetry on transmit and receive links between the master and slave. If the link delay between the master and slave is asymmetric, synchronization values will be inaccurate, and the inaccurate value will be half of the asymmetric link delay.

The transmission delay on a one-meter fiber is about 5 ns, so asymmetry for on a one-meter fiber causes a 2.5 ns time synchronization inaccuracy, and asymmetry on a 400-meter fiber causes a 1 us inaccuracy. On a practical network, it is difficult to accurately control the E2E optical fiber asymmetry on the entire network within 400 m. For TD-SCDMA and LTE-TDD which require a synchronization precision of +/-1.5 us, a 1 us inaccuracy is obviously intolerable. Therefore, in deployment of the IEEE 1588v2 time synchronous network, optical fiber asymmetry in the network must be seriously considered.

Solution

Currently, three solutions to fiber asymmetry are used in the world:

- Point-by-point measurement by instrument: During deployment and maintenance, a device's IEEE 1588v2 time synchronization precision is measured by an IEEE 1588v2 instrument point by point. Optical fiber asymmetry is speculated based on the measured precision, and compensation is made accordingly. Alternatively, an OTDR can be used to measure asymmetry of each pair of fibers, and compensation can be made. Because OTDR measurement causes service interruption and cannot directly display the results, an IEEE 1588v2 instrument is usually used.
- Bidirectional transmission over a single fiber: A device uses a bidirectional single-fiber optical module to support signal transmission and reception on a single optical fiber. This solution substantially resolves the asymmetry issue on dual-fiber links, and no measurement or compensation is required. In this solution, bidirectional link delay due to different wavelengths used for transmitting and receiving signals is generally small and can be automatically calculated and compensated by devices. According to the ITU-T G.652 Recommendation, the time inaccuracy on an FE bidirectional optical fiber is about 1.06 ns/km, and that on a GE bidirectional optical fiber is about 0.544 ns/km, both of which are far smaller than the inaccuracy on a normal dual-fiber link (2.5 ns/m). With the bidirectional single-fiber optical module used, onsite measurement is not required during deployment or maintenance.
- Automatic ring network measurement: A device uses a common dual-fiber optical module. During deployment, an instrument measures and compensates for asymmetry. During maintenance, asymmetry can be automatically measured on the ring network to avoid the necessity for onsite measurement upon fiber disconnection. This solution reduces the workload for maintaining fiber disconnection. This solution utilizes the BMC algorithm in IEEE 1588v2. When fiber disconnection occurs on the ring network, services are automatically switched to the standby link, and the time synchronization precision of the nodes on the ring network stays in the usable range. When the faulty fiber is re-connected, the devices first automatically calculate and report asymmetry on the new fiber link and then compensate for asymmetry on the NMS. After that, services are switched back to the active link.

Delay Compensation Mode

Delay compensation can be length compensation or time compensation.

The relationship between the length compensation and time compensation is as follows: the delay compensated for each 1 m twisted pair is 8 ns, and the delay compensated for each 1 m coaxial cable is 4.5 ns.

Delay Compensation for 1PPS+TOD Transmission on Cables

When a device uses the clock signals input through a 1PPS+TOD port, the cable connected to this port is measured in meters. When the clock signals are transmitted to a slave clock device through cables, a large delay occurs. The cable for transmitting 1PPS is unidirectional and

does not support automatic computation of compensation for bidirectional delay defined in 1588v2. Therefore, manual compensation is required. In practice, users must measure the cable length and then configure the delay compensation using the command line interface (CLI) or NMS.

NOTE

The external time port on the product is not a PTP port and does not support the IEEE 1588v2 protocol. The transmission delay cannot be measured automatically. Therefore, the transmission delay of the cable connecting to the external time port must be measured using a test instrument or computed based on the cable length.

32.6.4 1588v2 Network Application

Table 32-9

Scenario	Application Condition
Recommended: Synchronization Application (Network-wide 1588v2 and Synchronous Ethernet Deployment)	• Devices at the bearer network supports 1588v2 synchronization, time source is injected from the devices, and time synchronization is implemented using 1588v2. • Devices at the entire network support synchronous Ethernet (SyncE), and clock synchronization is implemented using SyncE.
Synchronization Application (Network-wide 1588v2 Deployment)	• Devices at the bearer network supports 1588v2 synchronization, time source is injected from the devices, and time synchronization is implemented using 1588v2. • Devices at the entire network do not support Synchronous Ethernet (SyncE), and clock synchronization is implemented using 1588v2.
Synchronization Application (Clock or Time Signal Injection from an OLT)	Devices at the bearer network do not support 1588v2 synchronization, and time source and cock source are injected from an OLT.
Time and Clock Synchronization Protection Application	A protection solution is provided on the bearer network and multiple time sources and clock sources are available.

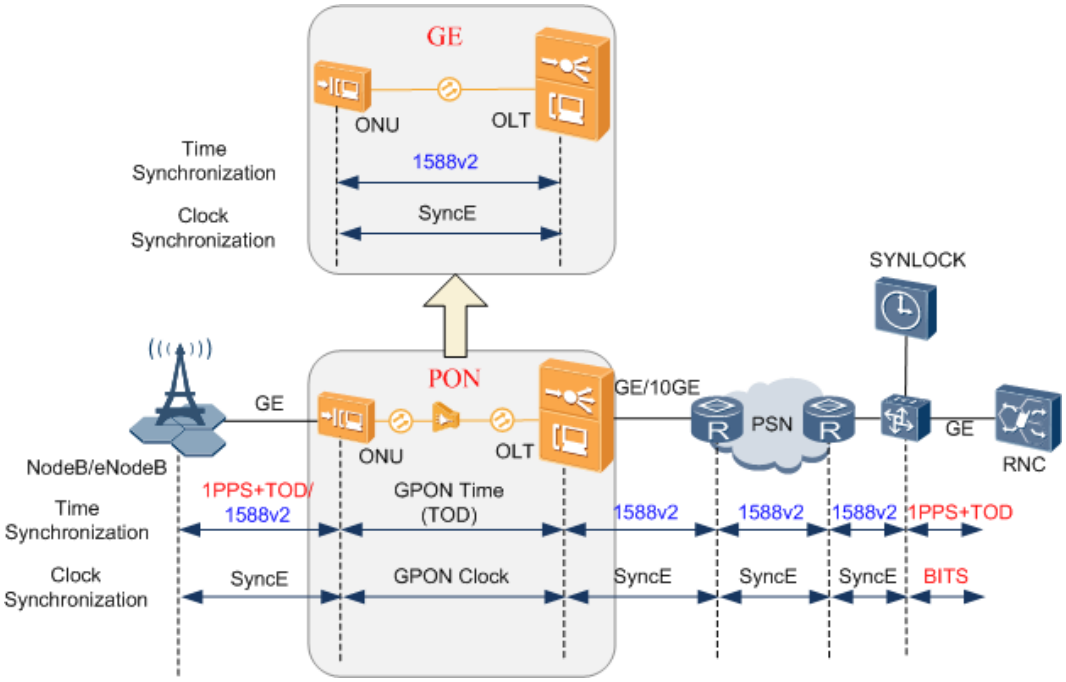
Recommended: Synchronization Application (Network-wide 1588v2 and Synchronous Ethernet Deployment)

Time synchronization must be implemented on 3G and 4G Long Term Evolution (LTE) base stations and can be provided by the bearer network. When bearer network devices support 1588v2 synchronization, time signals are injected to a bearer network device and be synchronized. When devices on the entire network support synchronous Ethernet (SyncE), SyncE is recommended for clock synchronization and 1588v2 for time synchronization, as shown in Figure 32-25.

NOTE

SyncE uses Ethernet bit streams to recover clock signals for Ethernet clock synchronization. The implementation mode is similar to that on an SDH/PDH network. SyncE uses a high-precision clock signals as the transit reference in the transmit direction. It restores and extracts the clock signals at the receive end. The physical layer transmits and receives the clock signals and is compatible with traditional networks.

Figure 32-25 Synchronization application (network-wide 1588v2 and SyncE deployment)



Time Synchronization

- Bearer network devices are used as a Ordinary Clock (OC) device. A bearer network device receives time signals through a 1PPS+TOD port and uses 1588v2 to synchronize time signals with other bearer network devices.
- OLTs and ONUs are used as a boundary clock (BC) device. An OLT receives 1588v2 time signals through a GE or 10GE link connected to a bearer network device. The OLT uses GPON or GE lines to transmit the time signals to an ONU in optical network terminal management and control interface (OMCI) mode. PTP packets are transmitted through GE lines to transmit time signals. Then the ONU uses a user-side GE link to transmit the signals to base stations.
- Base stations are used as a Ordinary Clock (OC) device. The base stations implement 1588v2 synchronization in different situations.
  - If a base station does not support 1588v2 synchronization, the base station uses a 1PPS+TOD port to receive time signals.
  - If a base station supports 1588v2 synchronization, the base station uses 1588v2 to synchronize time signals.

Clock Synchronization

- Bearer network devices use SyncE to synchronize clock signals with each other.

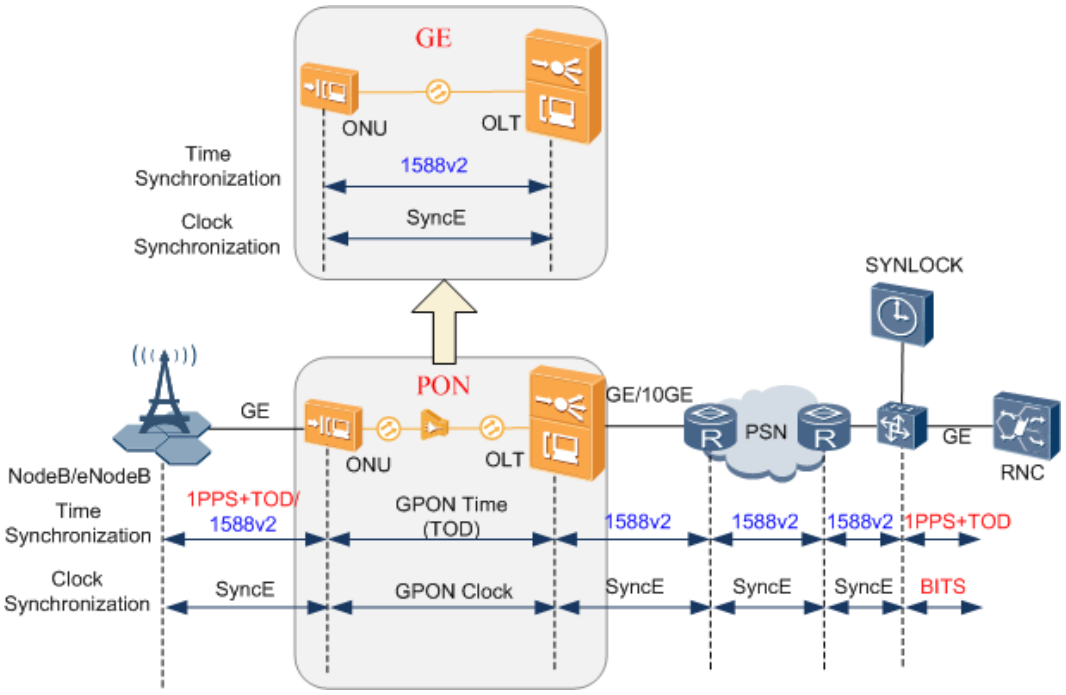
- An OLT receives clock signals from a GE or 10GE link. The OLT uses GPON lines to transmit the signals to an ONU in OMCI mode. Then the ONU uses a user-side GE link to transmit the signals to base stations.
- The base stations use SyncE to synchronize clock signals.

Synchronization Application (Network-wide 1588v2 Deployment)

Time synchronization must be implemented on 3G and 4G Long Term Evolution (LTE) base stations and can be provided by the bearer network. When bearer network devices support 1588v2 synchronization, time signals are injected to a bearer network device and be synchronized, as shown in Figure 32-26.

- Bearer network devices are used as a Ordinary Clock (OC) device. A bearer network device receives time signals through a 1PPS+TOD port and uses 1588v2 to synchronize time and clock signals with other bearer network devices.
- OLTs and ONUs are used as a boundary clock (BC) device. An OLT receives 1588v2 time signals through a GE or 10GE link connected to a bearer network device. The OLT uses GPON or GE lines to transmit the time signals to an ONU in optical network terminal management and control interface (OMCI) mode. PTP packets are transmitted through GE lines to transmit time signals. Then, the ONU uses a user-side GE link to transmit the time signals to base stations.
- Base stations are used as a Ordinary Clock (OC) device. The base stations implement 1588v2 synchronization in different situations.
  - If a base station does not support 1588v2 synchronization, the base station uses a 1PPS+TOD port to receive time signals, and use SyncE to synchronize clock signals.
  - If a base station supports 1588v2 synchronization, the base station uses 1588v2 to synchronize time and clock signals.

Figure 32-26 Synchronization application (network-wide 1588v2 deployment)

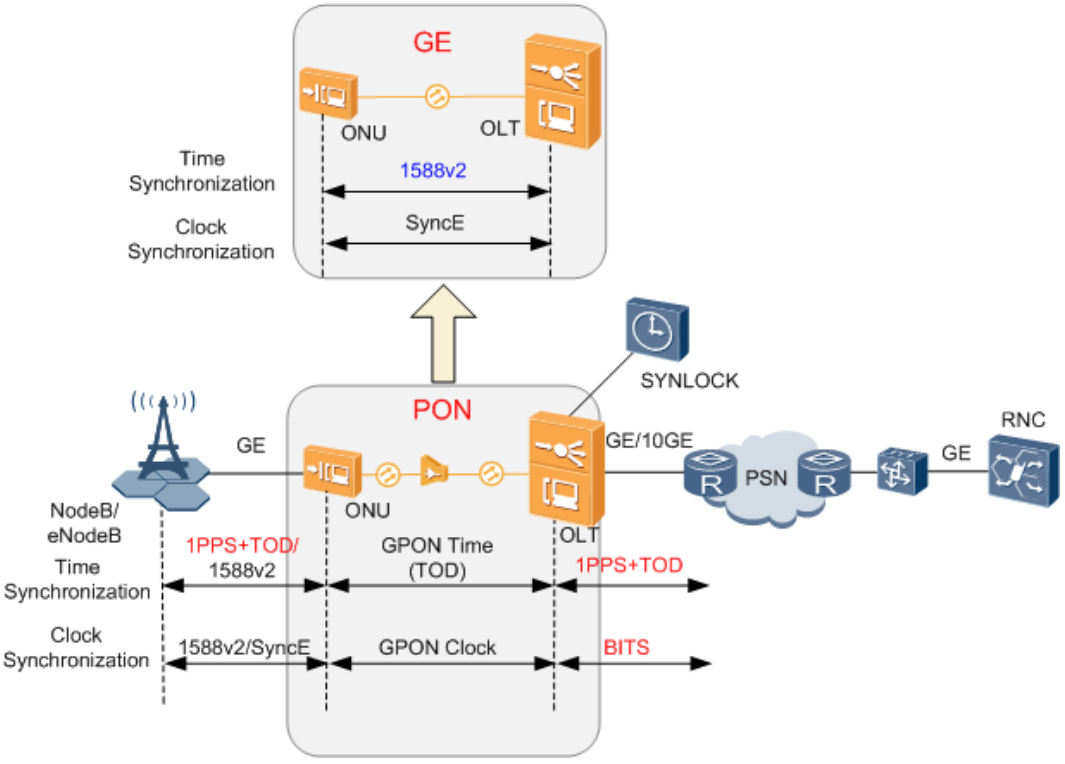


Synchronization Application (Clock or Time Signal Injection from an OLT)

3G and 4G long term evolution (LTE) base stations require time synchronization. When devices on the bearer network do not support 1588v2 synchronization, time signals can be injected from an OLT, as shown in Figure 32-27.

- Time signals are injected from an OLT using a 1PPS+TOD port. The BITS synchronization clock source is injected into the building integrated timing supply (BITS) port.
- OLTs and ONUs are used as a boundary clock (BC) device. The OLT uses GPON or GE lines to transmit the time signals to an ONU in optical network terminal management and control interface (OMCI) mode. PTP packets are transmitted through GE lines to transmit time signals. Then, the ONU uses a user-side GE link to transmit the time signals to base stations.
- Base stations are used as a Ordinary Clock (OC) device. The base stations implement 1588v2 synchronization in different situations.
  - If a base station does not support 1588v2 synchronization, the base station uses a 1PPS+TOD port to receive time signals, and use SyncE to synchronize clock signals.
  - If a base station supports 1588v2 synchronization, the base station uses 1588v2 to synchronize time and clock signals.

Figure 32-27 Synchronization application (clock or time signal injection from an OLT)



Time and Clock Synchronization Protection Application

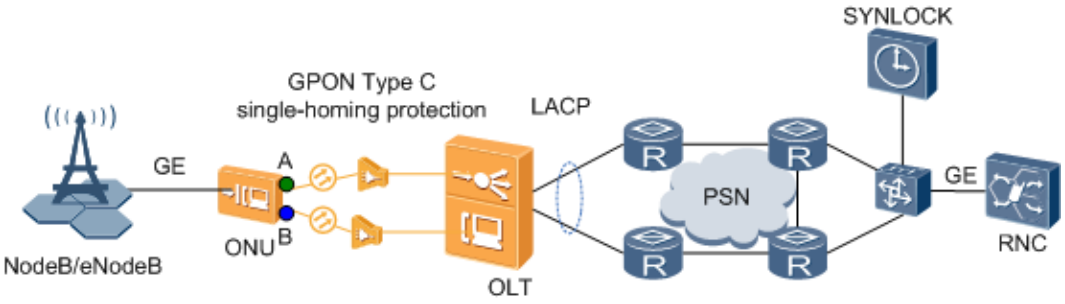
The time and clock synchronization protection feature improves network security.

As shown in Figure 32-28, GPON Type C single-homing protection is configured between the ONU and two OLTs, and LACP is configured between the OLTs and bearer network devices. When the working link becomes faulty, protection switching is triggered and data is switched to the standby link. This protection mechanism ensures uninterrupted data transmission.

As shown in Figure 32-28, port A is assumed to be a master port and port B is a slave port. If the active link is faulty, the clock source must also be switched to the standby link. That is, the link is switched from port A to port B. This prevents signals from being lost.

After the faulty link restores, the system will not immediately switch the clock source back to the original active link and the clock source of port B is preferentially traced.

Figure 32-28 GPON Type C dual-homing protection



NOTE

Other protection features can be deployed between the OLT and ONUs, or between the OLT and the bearer network. For details about the protection features, see feature description.

Table 32-10 Time and clock source selection modes on an ONU/OLT

Synchronization Type	Time and Clock Source Selection Modes on an ONU/OLT
Time synchronization	• Static selection based on priorities configured on an ONU/OLT Users must configure a unique priority for each external time source. The system selects the time source that has the highest priority and is in the normal state as the reference time. • Automatic selection based on best master clock (BMC) The BMC algorithm classifies clocks on a network into a master clock and slave clocks. Slave clocks trace the frequency or time of the master clock. When the network topology or the attribute of a clock source changes, the algorithm selects a new grandmaster clock so that clocks and times are synchronized on the entire network. For details about BMC, see BMC Algorithm.
Clock synchronization	• Static selection based on priorities configured on an ONU/OLT Users must configure a unique priority for each external clock source. The system selects the clock source that has the highest priority and is in the normal state as the reference clock. • Automatic selection based on Synchronization Status Message (SSM) levels The system obtains the clock quality level from received clock

Synchronization Type	Time and Clock Source Selection Modes on an ONU/OLT
	source signals (users must manually configure the quality level for a clock source whose SSM level cannot be obtained). Then the system automatically selects a clock source as the reference clock when the source meets the following three conditions: has the highest level; the level is not smaller than the lower SSM level; is in the normal state. If two external clock sources both have the highest quality level, the system selects the clock source that has a higher priority as the reference clock. For details about SSM, see Working Principle of Clock.

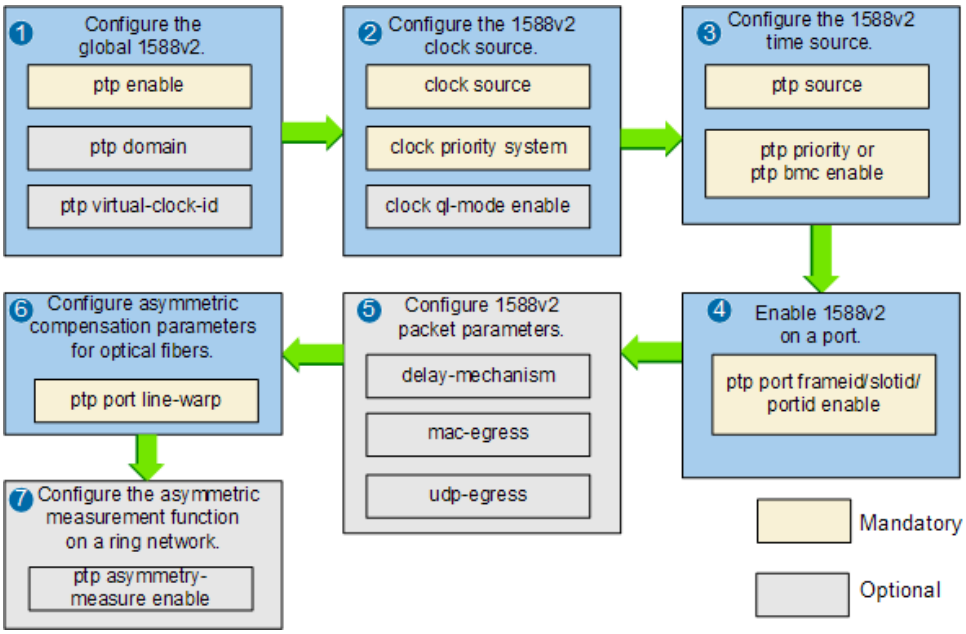
32.6.5 Configuring the 1588v2 Function

Universal mobile telecommunications system (UMTS) and long term evolution (LTE) base stations require frequency (clock) and phase (time) synchronization. This section describes how to configure 1588v2 to implement clock synchronization and sub-microsecond-level time synchronization.

Configuration Process

Figure 32-29 shows the process of configuring the 1588v2 function.

Figure 32-29 Process of configuring the 1588v2 function



Procedure

Configure the global 1588v2 function.

1. Run the **ptp enable** command to enable the 1588v2 function on a device.

2. **Optional:** Run the **ptp domain** command to specify a 1588v2 time domain.

   After the configuration, a time domain has a clock source. All devices in the time domain use the clock source. Ensure that the MDU time domain is the same as that of the OLT and upper-layer bearer network device.

3. **Optional:** Run the **ptp virtual-clock-id** command to configure the 1588v2 virtual clock ID of a device.

   The clock ID uniquely identifies a clock node in a 1588v2 domain.

   If this parameter is not configured, its value consists of the organizational unique identifier (OUI)+product ID+last four bytes of the MAC address. The OUI is 00-25-9e, and the product ID is 40. For example, if the MAC address of the system is 2017-5898-a025, the parameter value is 00259e405898a025.

Step 1 Configure a 1588v2 clock source. After the configuration, device clocks on the network are synchronized.

1. Run the **clock source 1588** command to configure a 1588v2 clock source.

   The clock source can be the synchronous Ethernet clock or 1588v2 clock recovered from a GE or 10GE line on the network side or the BITS clock.

   If the network supports the synchronous Ethernet clock, the clock source preferentially uses the synchronous Ethernet clock.

2. Run the **clock priority system** command to set the priority of the clock source. A device may trace multiple clock sources for synchronizing the clock. It selects the clock source with the highest priority as the clock source.

   p0 indicates the highest priority, and p9 indicates the lowest priority.

3. **Optional:** Configure the clock source selection mode when the protection mechanism is configured.

   A clock source can be statically or dynamically selected.

   - To statically select a clock source based on the local priority, perform Step 2.2 to set a local priority for a clock source. The device selects the clock source with the highest priority as the clock source.
   - To automatically select a clock source using the Synchronization Status Message (SSM):
     i. Run the **clock ql-mode enable** command to configure the status of the SSM source selection function.
     ii. Run **clock ql** command to configure the SSM quality level of the clock source. The device selects a clock source as the clock source when the clock source has the optimal quality level that is greater than or equal to the lowest SSM level of the clock input source and is running properly.

Step 2 Configure a 1588v2 time source. After the configuration, device time on the network is synchronized.

1. Run the **ptp source** command to configure a 1588v2 time source.

   The time source can be the 1588v2 time recovered from a GE or 10 GE line on the network side or BITS time. If the network does not support 1588v2 time synchronization, the time source can use the BITS time.

2. **Optional:** When the BITS time is used, run the **ptp bits** command to configure the attributes of a BITS clock source, such as the precision and class of the clock source, and the priority of the clock signal.

3. Select a time source when the protection mechanism is used.

   A time source can be statically or dynamically selected.

- To statically select a time source based on the local priority, run the **ptp priority** command to set the priority of the time source.
- To dynamically select a time source based on the best master clock (BMC), run the **ptp bmc enable** command to enable the automatic BMC source selection function.

Step 3 Enable 1588v2 on a port to transmit 1588v2 packets.

Run the **ptp port frameid/slotid/portid enable** command to enable 1588v2 on an Ethernet port.

Enable 1588v2 on the ports that are required to transmit 1588v2 packets, for example, the uplink port on an OLT, user-side port on an MDU and GE port connecting the OLT and MDU. You are not required to enable the function on the GPON port connecting the OLT and MDU.

Step 4 **Optional:** Configure 1588v2 parameters.

Set the timestamp mode of the 1588v2 packets, MAC encapsulation mode, and UDP encapsulation mode.

1. Run the **ptp port *frameid/slotid/portid* delay-mechanism { delay | pdelay }** command to configure the delay measurement mode for 1588v2 ports. The default mode is delay.

   The delay measurement mode on ports of all the devices must be the same.
2. Run the **ptp port *frameid/slotid/portid* clock-step { one-step | two-step }** command to configure the step mode for 1588v2 ports.

   By default, 1588v2 packets carry timestamps in one-step mode. The 1588v2 ports identify Follow_Up packets in two-step mode in the Rx direction for communicating with other products.
3. Run the **ptp port***frameid/slotid/portid***interval** { **announce***value* | **req***value* | **sync***value* } command to configure the interval at which 1588v2 packets are sent.
   - 1588v2 frequency synchronization:

     The peer-end master device must send Sync packets at an interval longer than or equal to 32 packets per second.
   - 1588v2 time synchronization:
     - The peer-end master device must send Sync packets at an interval longer than or equal to 1 packet per second.
     - The local-end device must send Delay_request or Pdelay_request packets at an interval longer than or equal to 1 packet per second.
4. **Optional:** Configure the encapsulation mode for 1588v2 packets.
   - Run the **ptp port***frameid/slotid/portid***mac-egress** [{ **destination-mac***destination-mac-value* | **vlan***vlan-id* [**priority***priority-value*] }] command to configure the MAC encapsulation mode for 1588v2 packets to be forwarded by the port. The packets are forwarded in Layer 2 mode.
     - When you specify **destination-mac**, the packets are forwarded in Layer 2 unicast mode.
     - When you do not specify **destination-mac** *destination-mac-value*, the packets are forwarded in Layer 2 multicast mode.
   - Run the **ptp port***frameid/slotid/portid***udp-egress** { [**destination-mac***destination-mac-value*] | [**source-ip***source-ip-addr* [**destination-ip***destination-ip-addr*][**dscp***dscp-value*][**vlan***vlan-id* [**priority***priority-value*]]] } command to configure the UDP encapsulation mode for 1588v2 packets to be forwarded on the port. The packets are forwarded in Layer 3 mode.

- When you specify **destination-ip** *destination-ip-addr* the packets are forwarded in Layer 3 unicast mode.
- When you do not specify **destination-ip** *destination-ip-addr*, the packets are forwarded in Layer 3 multicast mode.

Step 5 Configure asymmetric compensation parameters for optical fibers.

Run the **ptp port***frameid/slotid/portid***line-warp***linewarptype***linewarpdir***linewarpvalue* command to configure asymmetric compensation parameters for optical fibers.

Asymmetric optical fibers between two devices result in time difference for data transmission and reception. During site deployment, measure and calculate the time differences in the two directions and compensate the time on the devices by configuring cable transmission deviations.

Step 6 **Optional:** Configure the asymmetric measurement function on a ring network.

Run the **ptp asymmetry-measure enable** command to configure the asymmetric measurement function on a ring network.

This function utilizes the BMC algorithm in IEEE 1588v2. When fiber disconnection occurs on the ring network, services are automatically switched to the standby link, and the time synchronization precision of the nodes on the ring network stays in the usable range. When the faulty fiber is re-connected, the devices first automatically calculate and report asymmetry on the new fiber link and then compensate for asymmetry on the NMS. After that, services are switched back to the active link.

----End

Result

- Run the **display clock source** command to check whether the clock source in the system is 1588v2 and whether it is running properly. In normal cases, **State** is **successful**.
- Run the **display ptp source** command to check whether the clock source in the system is 1588v2 and whether it is running properly. In normal cases, **Lock state** is **successful**.

When clock tracing is abnormal, rectify the fault based on the handing suggestions in the reported alarm or event.

Example

An MDU connects to a 4G LTM eNodeB using a GE port. The MDU connects to the OLT using the GPON uplink port and then to the radio network controller (RNC) over the upper-layer network to carry 4G services over the access network. The 4G LTE eNodeB requires high precision time synchronization and the 1588v2 time is deployed on the network.

The parameters planned on the OLT are as follows:

- Time domain: 1
- Clock or time source input: GE uplink port 0/8/0
- Clock or time source output: GPON service port 0/3/1
- Packet encapsulation mode: MAC (default mode) (packets are forwarded in Layer 2 multicast mode)

The parameters planned on the MDU are as follows:

- Time domain: 1

- Clock or time source input: GPON uplink port 0/0/0
- Clock or time source output: GE service port 0/1/1
- Packet encapsulation mode: MAC (default mode) (packets are forwarded in Layer 2 multicast mode)

To configure the 1588v2 function with the preceding settings, do as follows:

Configure the OLT.

```
huawei(config)#ptp enable
huawei(config)#ptp domain 1
huawei(config)#clock source 0 1588
huawei(config)#clock priority system 0
huawei(config)#ptp source 0 0/8/0
huawei(config)#ptp port 0/8/0 enable
huawei(config)#ptp bmc enable
```

Configure the MDU.

```
huawei(config)#ptp enable
huawei(config)#ptp domain 1
huawei(config)#clock source 0 0/0/0
huawei(config)#clock priority system 0
huawei(config)#ptp source 0 0/0/0
huawei(config)#ptp port 0/1/1 enable
huawei(config)#ptp bmc enable
```

32.6.6 Configuring 1588v2-related Delay Compensation for Asymmetric Fibers

Prerequisites

A global positioning system (GPS) and clock signal tester, such as a Paragon tester, are available.

The 1588v2 precision of the upper-layer device meets usage requirements.

1588v2-related Delay Compensation Rules for Asymmetric Fibers

- After the first compensation on a network, if the 1588v2 precision meets clock precision requirements (1 us) of a base station, no more compensation is required for this base station.
- If the 1588v2 precision is required to improve (to be within 500 ns) for a base station, or the 1588v2 precision of some base stations fails to meet clock precision requirements after the compensation on the entire network, a second compensation is required for such base stations.
- If the capacity of a network needs to be expanded (for example, a ring is added), and the compensation on new base stations or sites must not change the original compensation data, separate compensation is performed on the new base stations or sites.

Context

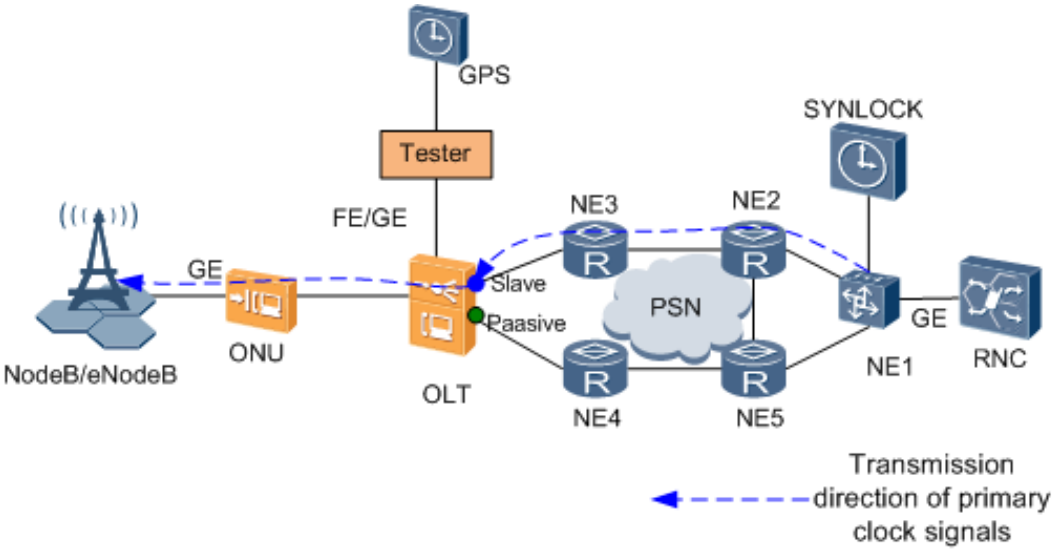
An OLT can connect to ONUs in PON or GE mode:

- In PON connections, the optical fibers between the OLT and ONUs are symmetric. Therefore, no asymmetric delay compensation is required for the optical fibers. In this case, you only need to measure the OLT and implement delay compensation on the optical fibers between the OLT and its upper-layer device.
- In GE connections, the optical fibers between the OLT and ONUs may be asymmetric. In this case, you need to measure the ONUs and implement delay compensation on the optical fibers not only between the OLT and its upper-layer device but also between the OLT and ONUs.

As shown in the following figure, the OLT is used as a test node.

- The OLT connects to a tester using an FE or GE port.
- The slave port on the OLT connects to NE 3 for receiving primary clock signals.
- The passive port on the OLT connects to NE 4 for receiving standby clock signals.

Figure 32-30 Delay compensation networking



Procedure

Perform delay compensation on the slave port of the OLT.

1. Calculate the difference between the data output by the FE or GE port after the OLT traces the 1588v2 clock and the data measured by the GPS. Record the difference as D1.
2. Check whether D1 is within ±20 ns. If it does, go to Step 2.1. If it does not, go to Step 1.3.
3. Run the **ptp port** *frameid/slotid/portid* **line-warp** *linewarptype linewarpdir linewarpvalue* command to set the compensation value of the slave port to d, which is equal to D1.
4. Calculate the difference between the output 1 PPS data after the OLT traces the 1588v2 clock and the data measured by the GPS. Record the difference as D2.
5. Check whether D2 is within ±20 ns. If it does, go to Step 2.1. If it does not, go to Step 1.3.

Step 1 Perform delay compensation on the passive port of the OLT.

1. Check whether the OLT provides a passive port. If it does, go to Step 3. If it does not, go to Step 2.2.
2. Run the **ptp port** *frameid/slotid/portid* **disable** command to disable 1588v2 on the original slave port.
3. Check whether the passive port has been switched to the slave port. If it does, go to Step 2.4.
4. Calculate the difference between the output 1 PPS data after the OLT traces the 1588v2 clock and the data measured by the GPS. Record the difference as D3.
5. Check whether D3 is within ±20 ns. If it does, go to Step 3. If it does not, go to Step 2.6.
6. Run the **ptp port** *frameid/slotid/portid* **line-warp** *linewarptype linewarpdir linewarpvalue* command to set the compensation value of the slave port to d, which is equal to D3.
7. Calculate the difference between the output 1 PPS data after the OLT traces the 1588v2 clock and the data measured by the GPS. Record the difference as D4.
8. Check whether D4 is within ±20 ns. If it does, go to Step 2.9. If it does not, go to Step 2.6.
9. Enable 1588v2 on the original slave port and check whether this port has been switched to the status queried in Step 2.1.

Step 2 Perform delay compensation on the uplink port of an ONU.

Connect a tester to the ONU and perform delay compensation on the optical fiber between the OLT and this ONU. For details, see Step 1.

----End

32.6.7 1588v2 Maintenance and Diagnosis

The 1588v2 clock or time source tracing status can be verified by running the following command:

- **display clock source**: used to check whether the system reference clock source is a 1588v2 clock source and whether the clock source is functional.
- **display ptp source**: used to check whether the system reference time source is a 1588v2 time source and whether the time source is functional.
- **display ptp info**: used to check whether the time locking status and 1588v2 running data are correct.

The following alarms and events reflect clock tracing faults. If any alarm or event is reported, rectify the fault.

Alarm	Name
0x2d31a005	The input signal of the 1588 external clock source is lost
0x2d31a001	The input signal of the one external clock source is lost
0x2d31a006	Clock is not in tracing mode

Event	Name
0x2d30a003	The system clock source switches to a new clock source
0x2d30a008	The input QL of clock source changes
0x2d31a007	Loss of ESMC
0x2d31a009	The 1588 clock source fails to be locked

32.6.8 1588v2 Reference Standards and Protocols

Document	Description
IEEE 1588-2008	Precision Clock Synchronization Protocol for Networked Measurement and Control Systems
ITU-T G.813	Timing requirements of SDH equipment slave clocks (SEC)
ITU-T G.823	The control of jitter and wander within digital networks which are based on the 2048 kbit/s hierarchy
ITU-T G.8260	Definitions and terminology for synchronization in packet networks
ITU-T G.8261	Timing and Synchronization aspects in Packet Networks
ITU-T G.8262	Timing characteristics of Synchronous Ethernet Equipment slave clock (EEC)

32.7 1588ACR

1588 adaptive clock recovery (ACR) enables the devices at the two ends of a packet switched network (PSN) to synchronize frequencies.

32.7.1 Why Is 1588 ACR Required

Feature Value

The 1588 adaptive clock recovery (ACR) feature proposed by IEEE implements frequency synchronization between the devices at the two ends of a packet switched network (PSN).

Compared with clock synchronization solutions used on old and new networks, the 1588 ACR feature implements clock synchronization with high accuracy and precision and low network deployment costs. The old networks include GSM mobile bearer networks, enterprise private line networks, and narrowband networks. The new networks include universal mobile telecommunications system (UMTS) networks and long term evolution (LTE) networks.

- The MAN network of an old network does not support clock synchronization by hop. Therefore, no accurate and precise clock synchronization solution is available.

- On a new network, synchronous Ethernet or 1588v2 is used to recover frequencies of convergence-layer devices by hop. This requires that the hardware in the entire network support synchronous Ethernet or 1588v2. The synchronous Ethernet and 1588v2 solutions feature clock synchronization with high accuracy and precision but high network deployment costs.

Function

1588 ACR synchronizes the frequencies of the devices at the two ends of a packet switched network (PSN). Specifically, a master device supporting 1588v2 encapsulates the local system clock data into 1588v2 packets. The PSN network transparently transmits the 1588v2 packets to a slave device. The slave device obtains timestamps from the received 1588v2 packets and recovers the clock data of the master device. The PSN network does not need to support clock synchronization, or the clock of the PSN network can be a third-party clock.

1588 ACR implements frequency synchronization but not time synchronization.

32.7.2 1588 ACR Basic Concepts

The basic concepts involved in the 1588 adaptive clock recovery (ACR) feature are the same as those of 1588v2. For details, see 32.6.2 1588v2 Basic Concepts of the 1588v2 feature.

1588 ACR Messages

ITU-T G.8265.1 defines seven types of 1588 adaptive clock recovery (ACR) messages. The Table 32-11 describes each type of 1588ACR messages.

Table 32-11 1588 ACR messages

Type	Classification	Message Transmit Rate (pps)	Function
Sync	Event message	1/16-128	Sync, Delay_Req, Follow_Up, and Delay_Resp messages carry timestamps for recovering clock frequencies.
Delay_Req			
Follow_Up	General message	1/16-128	
Delay_Resp			
Announce		1/16-8 Default: 1/2	Announce messages carry best master clock adaptive (BMCA) information, such as priorities and quality levels (QLs). The Announce messages are used by slave devices to select a clock source and by master devices to implement clock protection switchovers.
Signaling		None	Signaling messages are used for unicast negotiations. A slave device sends a master device a signaling message to request Sync or Announce messages. Then the master device sends a

Type	Classification	Message Transmit Rate (pps)	Function
			signaling message to respond to the slave device. The master and slave devices exchange Sync or Announce messages only after the negotiation between them is successful.
Management		N/A	Management messages are used for device management. ITU-T G.8265.1 has not defined how to use the management messages.

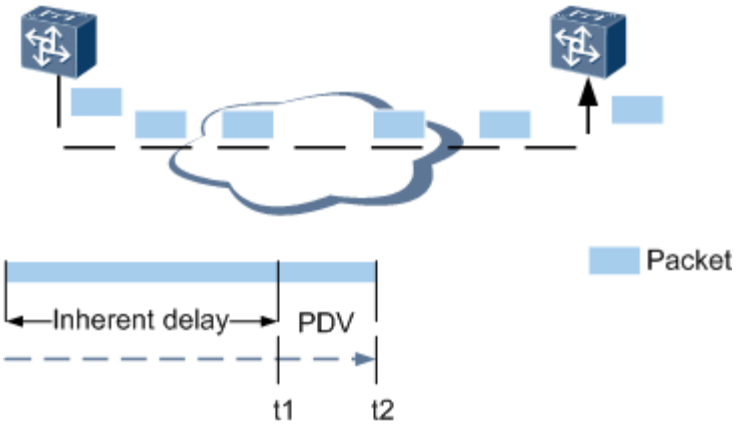
ITU-T G.8265.1 defines only transmit rate ranges for 1588v2 messages. Select a message transmit rate based on device hardware capability (crystal oscillation) and clock performance requirements (usage scenario and customer's requirements). A higher message transmit rate results in better clock recovery performance. The message transmit rate is configured on a slave device. After the configuration, the slave device sends the message transmit rate to the master device through a signaling request. Then the master device sends Sync or Announce messages to the slave device at the message transmit rate.

PDV

1588 adaptive clock recovery (ACR) is related to packet delay variation (PDV), which indicates the delay variation of the packet transmitted over a network. As shown in Figure 32-31, the packet forwarding delay is calculated using the following formula: Delay = Inherent delay + PDV.

- Inherent delay: the minimum delay when packets pass through a network. The inherent delay has a fixed value, which is the sum of the physical link (for example, optical fiber) delay and the fixed circuit delay inside a device.
- PDV: Because of queue scheduling by priority and clock domain translation, a device has to store packets in buffer registers, such as first in first out (FIFO) and random access memory (RAM), for a period of time (for example, several milliseconds) before sending them. As a result, the packets arrive at a destination after a delay, at any time from t1 to t2 in Figure 32-31. The delay is called PDV.

Figure 32-31 PDV diagram



1588 ACR Modes

Device Model

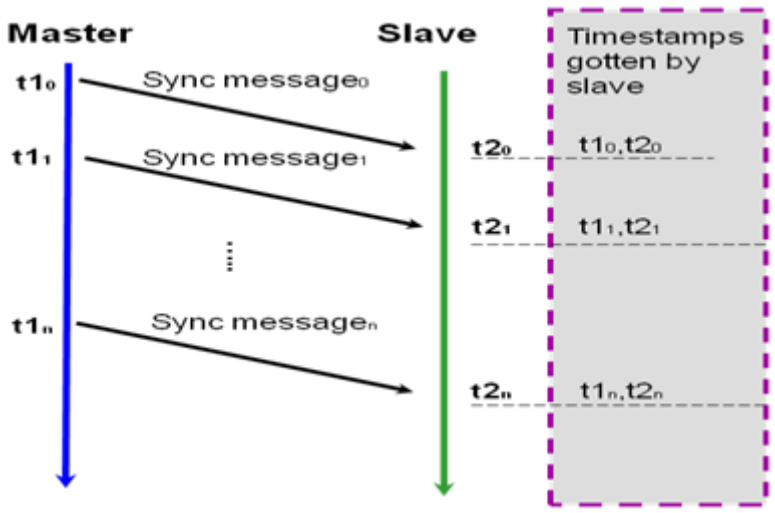
According to ITU-T G.8265.1, a slave device supporting 1588 adaptive clock recovery (ACR) supports only the slave-only ordinary clock (SOOC) model.

The 1588 ACR feature applies in end-to-end (E2E) deployment scenarios. A slave device needs to synchronize its clock frequency only with that of a master device. Therefore, the slave device does not need to support the boundary clock (BC) model.

One-way and Two-way

In one-way mode, the master device sends Sync messages to the slave device and the slave device does not send Delay_Req messages to the master device. The slave device recovers the frequency according to the timestamp contained in the Sync messages. For details, see Figure 32-32.

Figure 32-32 One-way mode



The master periodically sends Sync messages to the slave. If the slave frequency is synchronized to the master frequency, then the accumulative time errors within the same time periods are the same, as long as the path delay changes are neglected.

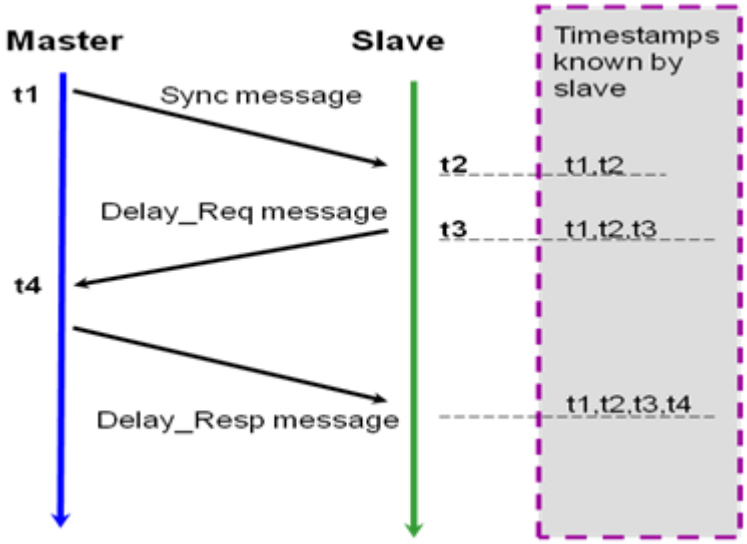
In other words, $t2_1-t2_0=t1_1-t1_0$, $t2_2-t2_1=t1_2-t1_1$, $t2_3-t2_2=t1_3-t1_2$,...$t2_n-t2_0=t1_n-t1_0$.

If $t2_n-t2_0$ is greater than $t1_n-t1_0$, then the slave frequency is higher than the master frequency, which means the slave frequency must be decreased. Reversely, the slave frequency must be increased.

The one-way mode implements basic principles of 1588ACR frequency synchronization and meets the requirements of 1588ACR frequency synchronization.

In two-way mode, the master device sends Sync messages to the slave device and the slave device sends Delay_Req messages to the master device. The bidirectional 1588 packet exchanging implements frequency and time synchronization. For details, see Figure 32-33.

Figure 32-33 Two-way mode



1. The master sends a Sync message at t1 and carries the t1 timestamp in the Sync message.
2. The slave receives the Sync message at t2, locally generates the t2 timestamp, and extracts the t1 timestamp from the Sync message.
3. The slave sends a Delay_Req message at t3 and locally generates the t3 timestamp.
4. The master receives the Delay_Req message at t4, locally generates the t4 timestamp, and sends the Delay_Req message with the t4 timestamp back to the slave.
5. The slave extracts the t4 timestamp from the Delay_Resp message after receiving it.

"Delayms" is the path delay in the direction from the master to the slave, "Delaysm" is the path delay in the direction from the slave to the master, and "Offset" is the time offset between the slave and master. All of them are variables.

t2-t1=Delayms+Offset

t4-t3=Delaysm-Offset

When Delayms = Delaysm, that is, when the transmit and receive links between the master and slave are symmetric, the following formula is satisfied:

Offset=[(t2-t1)-(t4-t3)]/2

The slave can calculate the time offset between itself and the master based on the t1, t2, t3, and t4 timestamps and then corrects its own time to get synchronized with the master.

Both the one-way and two-way modes implement frequency synchronization and the one-way mode uses fewer bandwidths. According to ITU-T G.8265.1, a master device must support both the modes, a slave device can support either of the modes, and a master device can support both one-way and two-way slave devices.

One-step and Two-step

The 1588 ACR mode can be one-step or two-step mode.

- In one-step mode, no Follow_Up messages are involved. A Sync message sent from a master device carries timestamp t1.
- In two-step mode, Follow_Up messages are used. After sending a Sync message, the master device generates a Follow-Up message to carry timestamp t1.

Figure 32-34 shows the one-step mode.

Figure 32-34 One-step mode

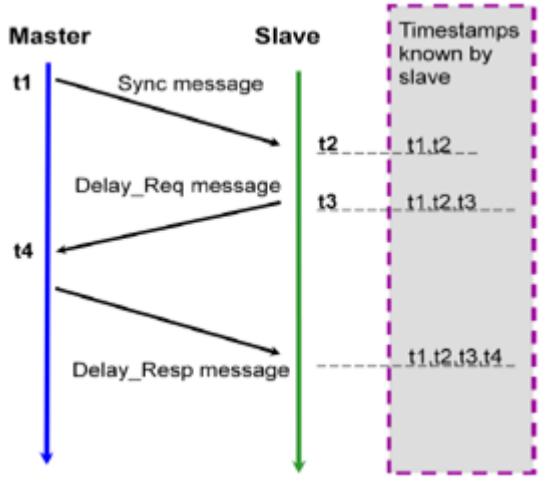
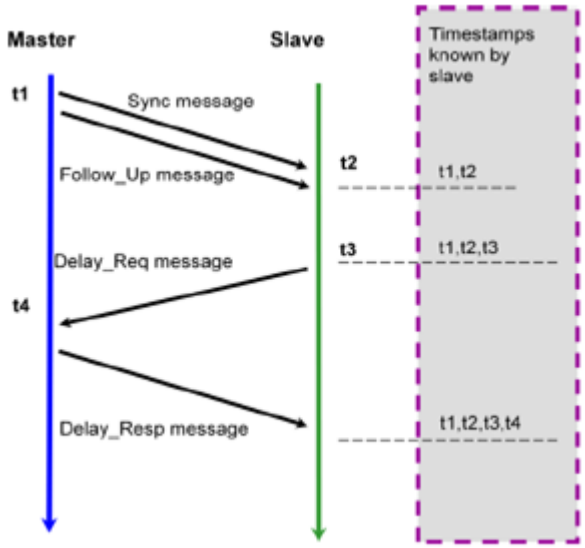


Figure 32-35 shows the two-step mode.

Figure 32-35 Two-step mode



The two-step mode is defined by IEEE 1588v2. In this mode, a Sync message corresponds to a Follow_Up message. When a Sync message passes the timestamp point, the master device hardware generates timestamp t1. The master device does not insert timestamp t1 into the Sync message, but inserts it into the Follow_Up message. The slave device extracts timestamp t1 from the received Follow_Up message. This reduces real-time requirements for the master and slave devices.

According to ITU-T G.8265.1, a 1588 ACR master device supports either or both of the one-step and two-step modes, and a 1588 ACR slave device automatically adapts to the mode carried in a Sync message sent from the master device. If the slave device connects to two master devices that use different 1588 ACR modes, the slave device must support the two modes.

Packet Encapsulation

According to ITU-T G.8265.1, 1588 ACR messages are encapsulated in UDP over IPv4 mode. This meets the requirements of appendix D *Transport of PTP over User Datagram Protocol over Internet Protocol Version 4* in the IEEE 1588v2 protocol.

1588 ACR messages are unicast UDP over IPv4 packets.

1588 ACR Clock Source Selection Algorithm

Based on the best master clock (BMC) algorithm of IEEE 1588v2, ITU-T G.8265.1 defines the best master clock adaptive (BMCA) algorithm for 1588 adaptive clock recovery (ACR). BMCA is simpler than BMC.

Table 32-12 lists BMCA parameters involved in 1588 ACR clock source selection.

Table 32-12 BMCA parameters involved in 1588 ACR clock source selection

Parameter	Description

Parameter	Description
Packet timing signal fail (PTSF)	Indicates that a Sync message is lost, an Announce message is lost, or a clock source is unavailable. The possible cause of an unavailable clock source is as follows: The packet delay variation (PDV) or the clock jitter of the master device is so large that the slave device cannot tolerate.
Quality level (QL)	Each master device carries its QL through an Announce message. The slave device preferentially traces the clock of the master device with the highest QL.
Priority	A slave device configures a local priority for each master device. If multiple master devices have the same QL, the slave device traces the clock of the master device with the highest priority.

If a slave device connects to multiple master devices and links are set up between the slave device and the master devices through unicast negotiations, the slave device receives the Announce and Sync messages sent from each master device. Then the slave device determines the best clock source as follows:

1. The slave device checks the PTSF of the clock of each master device. If PTSF occurs on a master device, the slave device does not trace the clock of the master device.
2. The slave device compares the QL of each master device and traces the clock of the master device with the highest QL.
3. If multiple master devices have the same QL, the slave device traces the clock of the master device with the highest priority.

Relationship Between 1588 ACR Clock Synchronization and Physical-Layer Clock Synchronization

The relationships between the 1588v2 clock synchronization and the physical-layer clock synchronization, including SDH synchronization and synchronous Ethernet, are as follows:

- The clock source of a 1588 ACR master device may be from an SDH or a synchronous Ethernet network.

  In this scenario, the master device must convert the synchronization status message (SSM) QL of the SDH or synchronous Ethernet network to the 1588 ACR clock class. In this manner, the clock sources of the master devices can be switched over for protection. In addition, a master device sends its clock QL to the slave device through Announce messages.
- The clock output from a 1588 ACR slave device can be used as the clock input to an SDH or synchronous Ethernet network.

  In this scenario, the 1588 ACR slave device recovers the clock frequency from 1588v2 messages and synchronizes the clock frequency with that of the master device. In addition, the slave device provides the recovered clock frequency to the downstream network or base station through the SDH or synchronous Ethernet network. In this case, the slave device must convert the 1588 ACR clock class to the SSM QL of the SDH and

synchronous Ethernet network. In this manner, the clock sources of the slave devices can be switched over for protection.

ITU-T G.8265.1 defines the mapping between the 1588 ACR clock class and the SSM QL of the SDH or synchronous Ethernet network, as shown in the Table 32-13.

Table 32-13 Mapping between the 1588 ACR clock class and the SSM QL of the SDH or synchronous Ethernet network

SSM QL	ITU-T G.781 Option	1588 ACR Clock Class
0010	QL-PRC	84
0100	QL-SSU-A	90
1000	QL-SSU-B	96
1011	QL-SEC	104
1111	QL-DNU	110

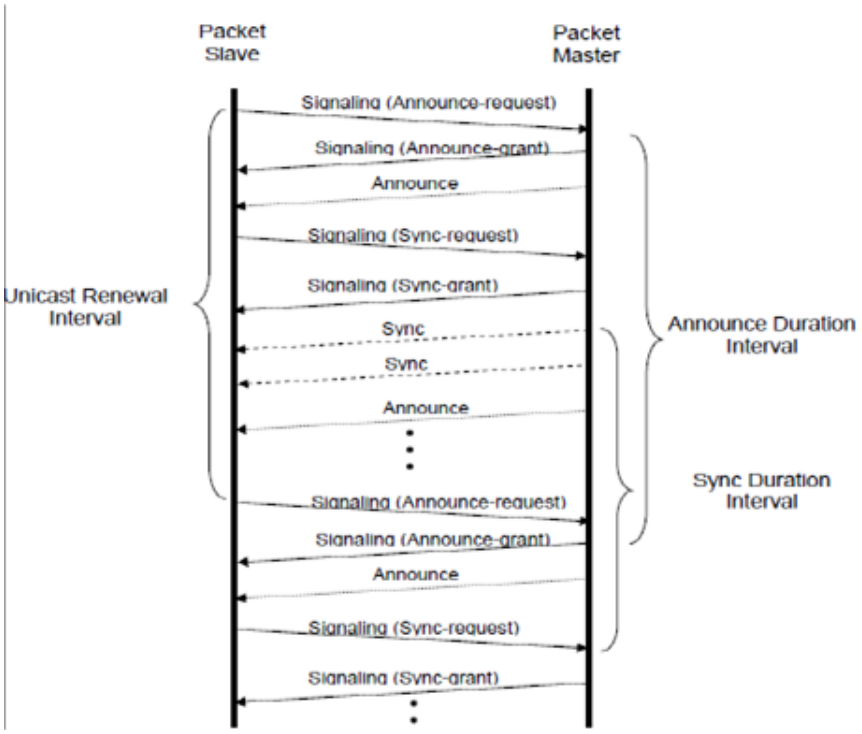
32.7.3 1588 ACR Principles

1588 ACR Unicast Negotiation

ITU-T G.8265.1 defines that a slave device communicates with a master device through a unicast negotiation. Specifically, the slave device initiates a request to the master device. The master device sends 1588v2 packets to the slave device only after authenticating the slave device.

A unicast negotiation is performed through signaling message that carry request data. The request data includes the requested message type, message transmit rate, and duration. The duration defines the valid period of a request. The master device sends the messages requested by the slave device only within the duration. After the duration elapses, the master device stops sending the messages to the slave device. The master device will start sending messages again only after the slave device initiates a new request. Figure 32-36 shows the process of a 1588 adaptive clock request (ACR) unicast negotiation.

Figure 32-36 Process of a 1588 ACR unicast negotiation



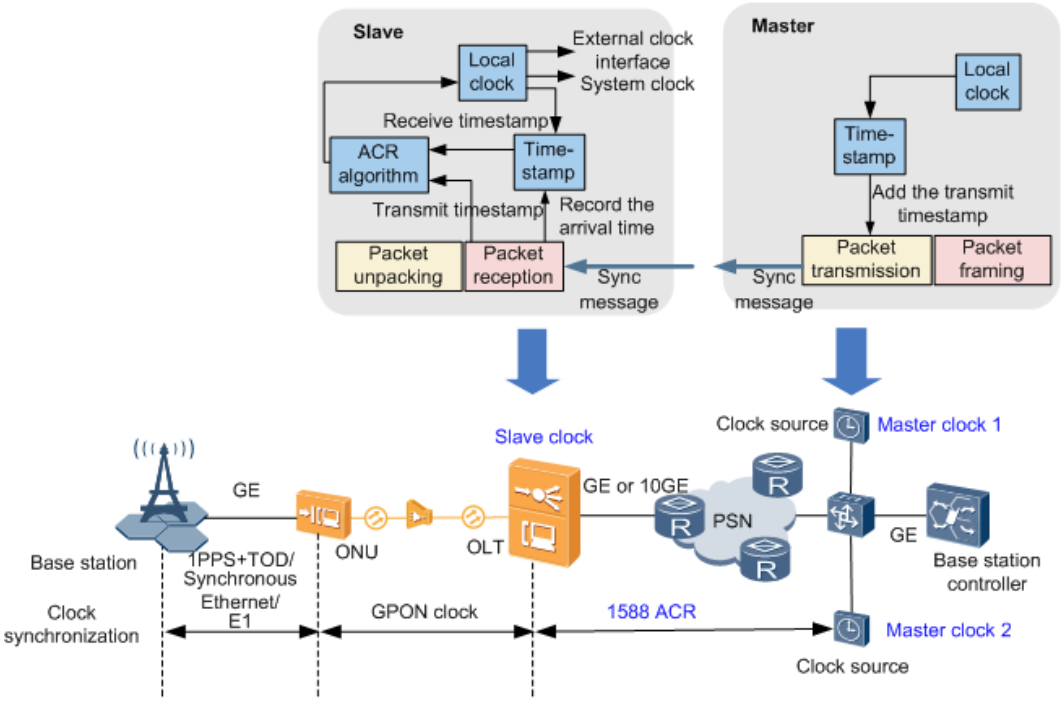
1. The slave device sends the master device a signaling message to request Announce messages. The signaling message contains the type, message transmit rate, and duration of the requested Announce messages.
2. The master device sends a signaling message to respond to the slave device, notifying the slave device that the request is permitted or rejected.
   - If the request is permitted, the master device sends Announce messages to the slave device at the message transmit rate requested by the slave device.
   - If the request is rejected, the master device does not send Announce messages to the slave device.
3. The slave device sends the master device a signaling message to request Sync messages. The signaling message contains the type, message transmit rate, and duration of the requested Sync messages.
4. The master device sends a signaling message to respond to the slave device, notifying the slave device that the request is permitted or rejected.
   - If the request is permitted, the master device sends Sync messages to the slave device at the message transmit rate requested by the slave device.
   - If the request is rejected, the master device does not send Sync messages to the slave device.

ITU-T G.8265.1 defines that the duration of each message can be configured on the slave device. The duration ranges from 60s to 1000s, with 300s by default. Before duration elapses, the slave device determines whether to initiate a new request. To ensure message continuity, the slave device initiates a new request after the duration elapses and before the master device stops sending messages.

1588 ACR Principles

1588 adaptive clock recovery (ACR) uses the ACR algorithm to recover the clock frequency of the transmit end from the transmit and receive timestamps carried in 1588v2 packets.

Figure 32-37 1588 ACR principles



The 1588 ACR process is as follows:

1. The slave device uses the 1588 ACR clock source selection algorithm to determine the best master clock from master devices 1 and 2 for tracing.
2. The master device providing the best master clock adds a transmit timestamp that contains the local system time to a 1588v2 packet.
3. The packet switched network (PSN) transparently transmits the 1588v2 packet to the slave device.
4. The slave device extracts the transmit timestamp from the received 1588v2 packet and adds the receive timestamp to the 1588v2 packet.
5. The slave device uses the ACR algorithm to process the transmit and receive timestamps contained in the 1588v2 packets that are received within the specified duration. Then the slave device recovers the clock frequency of the master device.

32.7.4 1588 ACR Deployment Requirements

When planning and deploying 1588 adaptive clock recovery (ACR), minimize the packet delay variation (PDV) for a stable running environment.

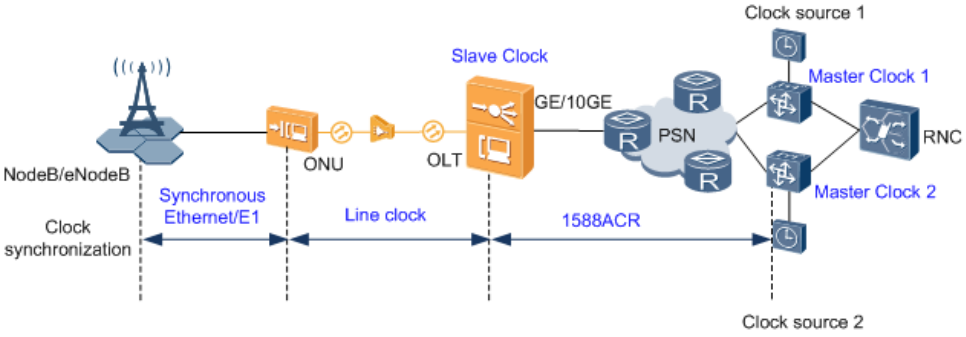
- Clock source protection: Deploy at least two master devices to implement clock source protection switchover.

- Locations of master and slave devices: Deploy the master and slave devices on the edge nodes connected to the intermediate network. This reduces the PDV caused by the intermediate network.
- Scale of an intermediate network: Adjust the routes of 1588 ACR messages and the locations of master and slave devices to minimize the number of forwarding hops for the 1588 ACR messages.
- Type of an intermediate network: Test and collect the PDV data on the live network to check whether 1588 ACR can be deployed on the network.
- Network traffic: Ensure that the long-term average traffic of the network is less than 80% of the maximum traffic that can be afforded by the network. Short-term network congestion and service interruptions are allowed according to ITU-T G.8265.1.
- Quality of service (QoS) priority: Configure a high priority for 1588 ACR messages to reduce the PDV caused by packet forwarding.
- Packet loss ratio: Ensure that the packet loss ratio is not greater than 0.5%. This prevents timing signal failure.
- Message transmit rate:
  - A higher transmit rate of Sync messages results in better clock recovery performance.
  - A higher transmit rate of Announce messages results in faster fault detection and protection switchover.
  - If there are many slave devices, a higher message transmit rate results in a higher bandwidth. In this case, the one-step mode is recommended. If the bandwidth is still high after the one-step mode is used, decrease the message transmit rate.

32.7.5 1588 ACR Networking

Figure 32-38 shows the 1588 adaptive clock recovery (ACR) networking.

Figure 32-38 1588 ACR networking



In the preceding figure:

- Clock sources 1 and 2 inject clock signals to the primary and secondary master devices, respectively. The clock signals are transmitted through 1588v2 or synchronous Ethernet. The convergence layer is a third-party packet switched network (PSN) or an intermediate network that does not support 1588v2 or synchronous Ethernet. 1588 ACR is deployed on the convergence layer to ensure clock precision.

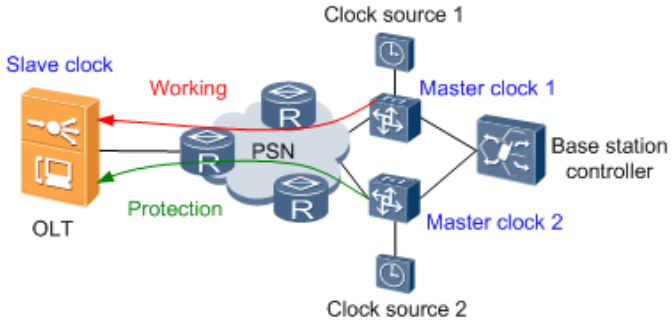
- The optical line terminal (OLT), a slave device, supports the negotiation with two master devices. According to the 1588 ACR clock source selection algorithm, the OLT traces the clock of the master device that provides better clock signals.
- The master devices convert 1588v2 or synchronous Ethernet clock packets to timestamps and encapsulate the timestamps into 1588 ACR messages. The convergence layer transparently transmits the 1588 ACR messages to the OLT. The OLT recovers the clock frequency from the received 1588 ACR messages and sends the clock frequency to the optical network unit (ONU) through line clock synchronization.
- The ONU sends the clock frequency to the base station through synchronous Ethernet, or E1 line clock synchronization.

The OLT is located at the end of a mobile bearer network or an access network. If 1588 ACR is enabled, the OLT can be used only as a slave device.

1588 ACR Protection Networking

If a clock source is faulty, services are interrupted. Therefore, the clock source must be protected. For the 1588 ACR feature, deploy a secondary master device to provide a secondary clock source. Figure 32-39 shows the 1588 ACR protection networking.

Figure 32-39 1588 ACR protection networking



In the preceding networking, if the primary master device is faulty, the slave device automatically traces the clock of the secondary master device. After the primary master device recovers, the slave device automatically traces the clock of the primary master device.

32.7.6 Configuring 1588 ACR

This section describes how to configure 1588 adaptive clock recovery (ACR) to implement the end to end (E2E) frequency synchronization across a third-party network.

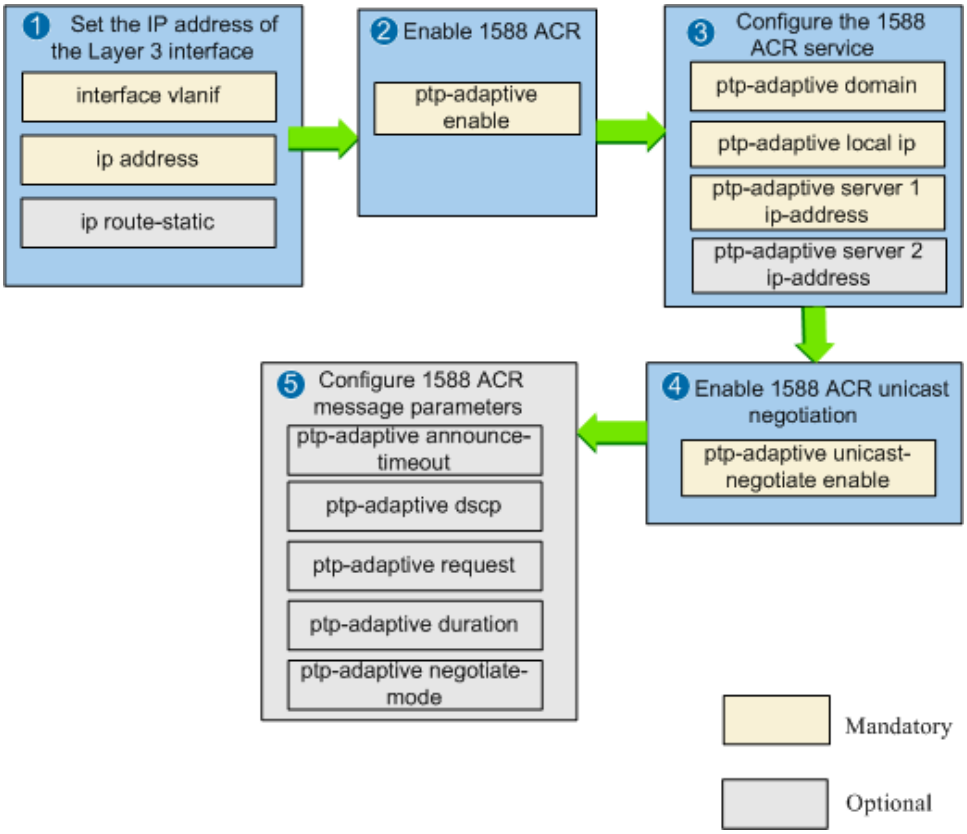
Prerequisites

- The optical line terminal (OLT) has been installed and deployed.
- The master device is functional.
- The route on the third-party network is reachable.

Configuration Process

Figure 32-40 shows the 1588 ACR configuration process.

Figure 32-40 1588 ACR configuration process



Procedure

Set the IP address of the Layer 3 interface.

1. Run the **interface vlanif** command to create a VLAN interface.
2. Run the **ip address** command to set the IP address of the VLAN interface.

NOTE

Ensure that the IP address of the VLAN interface and that of the master device port are in the same network segment.

3. **Optional:** Run the **ip route-static** command to configure a route from the local IP address of the master device to the IP address of the VLAN Layer 3 interface on the slave device.

   Perform this step if the local IP address of the master device port and that of the VLAN Layer 3 interface on the slave device are in different network segments.

Step 1 Run the **ptp-adaptive enable** command to enable 1588 ACR.

Step 2 Configure the 1588 ACR service.

1. Run the **ptp-adaptive domain** command to configure the time domain for the master and slave devices.

NOTE

Clock synchronization can be performed on the slave device only if the master and slave devices are in the same time domain.

2. Run the **ptp-adaptive local-ip** command to set the local IP address of the slave device.

1588 ACR messages are Layer 3 unicast IP packets. Therefore, a fixed IP address is required between the master and slave devices. This IP address is generally a loopback address or the IP address of an outbound port. In this step, the local IP address is the IP address of the VLAN interface set in Step 1.2.

3. Run the **ptp-adaptive server 1 ip-address** command to set the IP address of master device 1.

NOTE

In this step, the IP address is the local IP address of the master device.

4. **Optional:** Run the **ptp-adaptive server 2 ip-address** command to set the IP address of master device 2.

   Perform this step if a protection switchover is required for master devices.

NOTE

If a slave device connects to two master devices, the slave device traces the clock of the master device with a higher quality level and priority based on the 1588 ACR clock source selection algorithm. If the primary master device is faulty, the slave device automatically traces the clock of the secondary master device based on the 1588 ACR clock source selection algorithm.

Step 3 Run the **ptp-adaptive unicast-negotiate enable** command to enable 1588 ACR unicast negotiation.

Step 4 **Optional:** Configure 1588 ACR message parameters.

- Run the **ptp-adaptive announce-timeout** command to configure the maximum number of times that Announce messages are not received within a packet transmit period. The default number is 3.
- Run the **ptp-adaptive dscp** command to configure the priority for processing forwarded 1588 ACR messages. The default priority is 56.

NOTE

This parameter determines the 1588 ACR message forwarding priority. A higher priority results in a smaller network jitter for the 1588 ACR messages. Accordingly, the recovered clock performance is better. Otherwise, the recovered clock performance is not so good.

- Run the **ptp-adaptive request interval** command to configure the interval requested by the slave device at which the master device sends Announce, Sync, or Delay_Req messages to the slave device.

NOTE

- Default interval for Announce messages: 11
- Default interval for Sync messages: 3
- Default interval for Delay_Req messages: 3

The minimum message transmit rate required by the frequency recovered through 1588 ACR varies according to network condition. A larger packet transmit rate results in better recovered frequency performance. However, this requires more system resources and bandwidths. Therefore, configure a proper value to balance the frequency performance, network resources, and bandwidths.

- Run the **ptp-adaptive duration** command to configure the interval at which a re-negotiation is performed between the master and slave devices. The default time is 300s.

NOTE

The duration for each type of packet is separately configured. Therefore, before a duration elapses, the slave device determines whether to initiate a new request based on site requirements. To ensure message continuity, the slave device initiates a new request after the duration elapses and before the master device stops sending messages.

- Run the **ptp-adaptive negotiate-mode** command to configure the 1588 ACR negotiation mode to one-way or two-way. The default mode is one-way.

NOTE
The frequency recovered in two-way mode is better than that recovered in one-way mode, but this mode requires more bandwidths.

----End

Result

After the **display ptp-adaptive all** command is executed, **Ptp adaptive state** is **Enable**, **Current state** is **Slave**, and **Nego-state** is **Success**.

If an error occurs in clock tracing, rectify the fault based on the reported clock alarm or event.

Example

The following configurations are used as an example to configure 1588 ACR: An OLT connects to a third-party network through GE upstream transmission and then to a radio network controller (RNC). This implements wireless service carried over a base station network. The third-party network does not support 1588v2. Therefore, 1588 ACR is deployed on the network.

Data plan on the OLT:

- Time domain: 1
- VLAN ID: 2
- IP address of the VLAN interface: 10.10.11.2
- IP address of master device 1: 10.10.11.10

OLT configuration

```
huawei(config)#interface vlanif 2
huawei(config-if-vlanif2)#ip address 10.10.11.2 24
huawei(config-if-vlanif2)#quit
huawei(config)#ptp-adaptive enable
huawei(config)#ptp-adaptive domain 1
huawei(config)#ptp-adaptive local-ip 10.10.11.2
huawei(config)#ptp-adaptive server 1 ip-address 10.10.11.10
huawei(config)#ptp-adaptive unicast-negotiate enable
```

32.7.7 1588 ACR Maintenance and Diagnosis

1. If any of the following alarms or events is reported, rectify faults using the alarm or event troubleshooting methods.

Alarm	Name
0x2d31a005	The 1588 input signal of the external clock source is lost
0x2d31a001	The input signal of the one external clock source is lost
0x2d31a006	Clock is not in tracing mode

Event	Name
0x2d30a003	The system clock source switches to a new clock source
0x2d30a008	The input QL of clock source changes
0x2d31a007	Loss of ESMC
0x2d31a009	The 1588 clock source fails to be locked

2. Run the **display ptp-adaptive all** command to query the 1588 ACR running status. Rectify faults, if any.
3. Run the **display ptp-adaptive server** command to query the statistics of the collected transmit and receive 1588 ACR messages. Rectify faults, if any.
4. Run the **display ptp-adaptive all** command to query 1588 ACR configurations. Ensure that 1588 ACR is enabled, the IP address of the 1588 ACR server is correct, and 1588 ACR attributes are correct.
5. Check clock hardware status. Ensure that the upstream interface board, physical ports and links, and the clock daughter board are functional. Specifically, no board or port alarm is reported. A clock daughter board fault can be queried through alarms or commands.
6. Run the **display ptp-adaptive all** command to check the status and configuration of the Layer 3 interface.
7. Run the **display ip routing-table** command to check whether the master and slave routes are reachable.

32.7.8 1588 ACR Standard and Protocol Compliance

Standard or Protocol	Description
IEEE 1588-2008	Precision Clock Synchronization Protocol for Networked Measurement and Control Systems
ITU-T G.8260	Definitions and terminology for synchronization in packet networks
ITU-T G.8261	Timing and Synchronization aspects in Packet Networks
ITU-T G.8262	Timing characteristics of Synchronous Ethernet Equipment slave clock (EEC)
ITU-T G.8265	Architecture and requirements for packet based frequency delivery
ITU-T G.8265.1	Precision time protocol telecom profile for frequency synchronization

32.8 NTP

The Network Time Protocol (NTP) is used to synchronize the time between the distributed time server and the client.

32.8.1 NTP Introduction

Definition

The Network Time Protocol (NTP) is an application layer protocol in the TCP/IP protocol suite. NTP is used to synchronize the time between the distributed time server and the client. The implementation of NTP is based on IP and UDP.NTP involves the Time Protocol and the ICMP Timestamp Message, with special design on accuracy and robustness.

Purpose

NTP defines the accurate time in an entire network. Because the network topology is complicated, the clock synchronization among all the devices in the entire network becomes more critical.

The objective of NTP is to synchronize the clocks of all the devices on a network which have clocks. This helps to keep time consistency among all the devices in the network. Therefore, the device can offer various applications based on the clock synchronization. For example, before analyzing logs collected from different devices for network management, you must ensure that the time on the logs is from synchronized clocks.

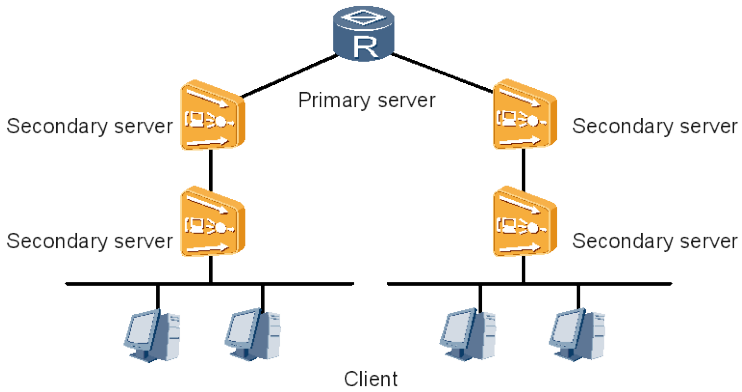
The MA5600T/MA5603T/MA5608T supports the NTP feature to guarantee that the clocks of all the devices in a network are consistent.

32.8.2 NTP Principle

NTP Network Architecture

As shown in the followed figure, the networking of NTP is composed of primary time server, secondary time server, clients, and interconnecting transmission paths.

Figure 32-41 Network Architecture of NTP



- A primary time server is directly synchronized with a primary reference source, which is usually a radio clock or Global Positioning System (GPS).
- A secondary time server synchronizes its clock with the clock of the primary time server on the network or other secondary time servers, and transmits the time information to clients on the network through NTP.

Under normal circumstances, primary and secondary time servers in the synchronization subnet assume a hierarchical-master-slave structure, with the primary server at the root and the secondary server at successive stratums toward the leaf node. The higher the stratum level is, the less accurate the clock will be.

NTP Operating Mode

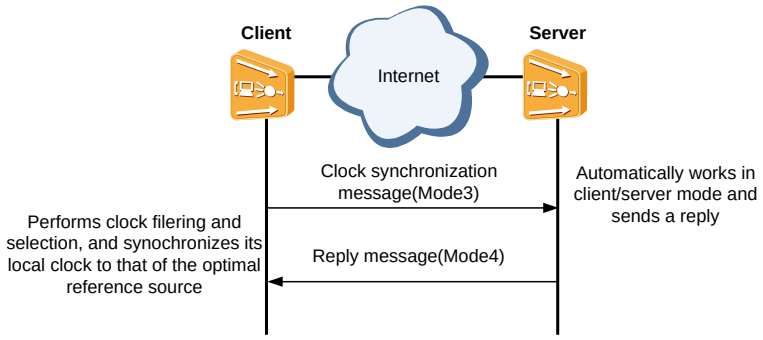
In actual application, you need to select a proper NTP operating mode based on the network deployment to meet various clock synchronization requirements. The operating modes of NTP are classified into Unicast Client/Server Mode, peer mode, broadcast mode, multicast mode and manycast mode.

Unicast Client/Server Mode

- The host that functions as a client sends packets to the server periodically. The value of the Mode field in a packet is set to 3. This indicates that the packet is sent by a client, without considering whether the server is reachable and which stratum the server is on. Usually, the host operating in client mode is a workstation on a specified network, which synchronizes its clocks with the clock on the server but does not alter the clock of the server.
- The host that functions as a server receives the packets from the client and sends response packets. The value of the Mode field in a response packet is set to 4. This indicates that the packet is sent by a server. Usually, the host operating in server mode is a time server on a network, which provides synchronization information for the clients but does not alter its own clock.

During and after the restart, the host operating in client mode periodically sends NTP request messages to the host operating in server mode. After receiving the NTP request message, the server swaps the position of destination IP address and source IP address, and the source port number and destination port number, fills in the necessary information, and sends the message to the client. The server does not need to retain state information when the client sends the request message. The client freely adjusts the interval for sending NTP request messages according to the local conditions.

Figure 32-42 Unicast Client/Server Mode



Peer Mode

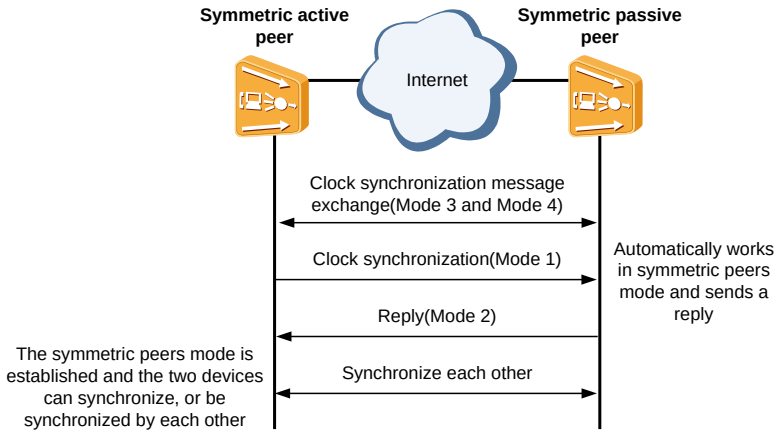
In this mode, the active peer and the passive peer can be synchronized with each other. To be specific, the higher stratum (lower level) peer is synchronized with the lower stratum (higher level) peer. The active and passive peers firstly exchange NTP packets whose values of Mode fields are 3 (sent by the client) and NTP packets whose values are 4 (sent by the server).

- Active peer: A host that functions as an active peer sends packets periodically. The value of the Mode field in a packet is set to 1. This indicates that the packet is sent by an active peer, without considering whether its peer is reachable and which stratum its peer is on. The active peer can provide time information about the local clock for its peer, or synchronize the time information about the local clock based on that of the peer clock.
- Passive peer: A host that functions as a passive peer receives packets from the active peer and sends response packets. The value of the Mode field in a response packet is set to 2. This indicates that the packer is sent by a passive peer. The passive peer can provide time information about the local clock for its peer, or synchronize the time information about the local clock based on that of the peer clock.
- Prerequisites for a host to function as a passive peer: The packets received by the local host are sent by an active peer. The number of the stratum that the active peer is on must be less than or equal to the number of the stratum that the local host is on. In addition, the routes between the local host and the active peer must be reachable.

NOTE

The host operating in passive mode is at the lower stratum in the synchronization subnet. You do not need to obtain information about the peer in advance because the connection between peers is not set up and status variables are not configured unless the passive host receives NTP messages from the peer.

Figure 32-43 Peer mode



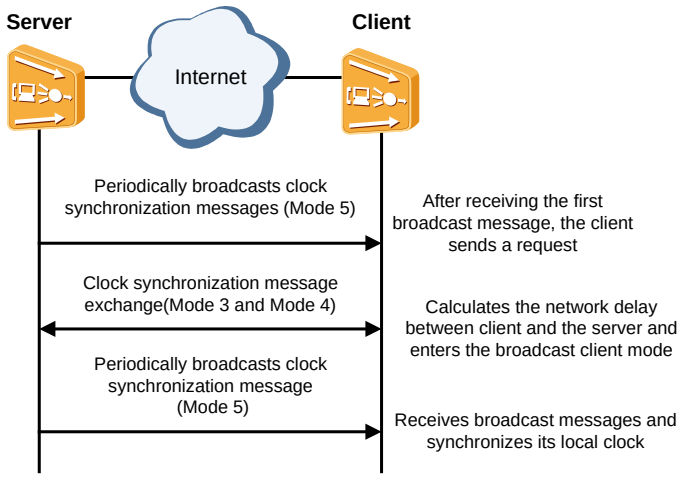
Broadcast Mode

- A host that runs in broadcast mode sends clock synchronization packets to the broadcast address 255.255.255.255 periodically. The value of the Mode field in a packet is set to 5. This indicates that the packet is sent by a host that runs in broadcast mode, without considering whether its peer is reachable and which stratum its peer is on. The host running in broadcast mode is usually a time server running high-speed broadcast media on the network, which provides synchronization information for all of its peers but does not alter the clock of its own.

- The client listens to the broadcast packets sent from the server. When the client receives the first broadcast packet, the client and server exchange NTP packets whose values of Mode fields are 3 (sent by the client) and the NTP packets whose values of Mode fields are 4 (sent by the server). In this process, the client enables the server/client mode for a short time to exchange information with the remote server. This allows the client to obtain the network delay between the client and the server. Then, the client returns the broadcast mode, and continues to sense the incoming broadcast packets to synchronize the local clock.

The broadcast mode is applied to the high speed network that has multiple workstations and does not require high accuracy. In a typical scenario, one or more time servers on the network periodically send broadcast packets to the workstations. The delay of packet transmission in a LAN is at the milliseconds level.

Figure 32-44 Broadcast mode

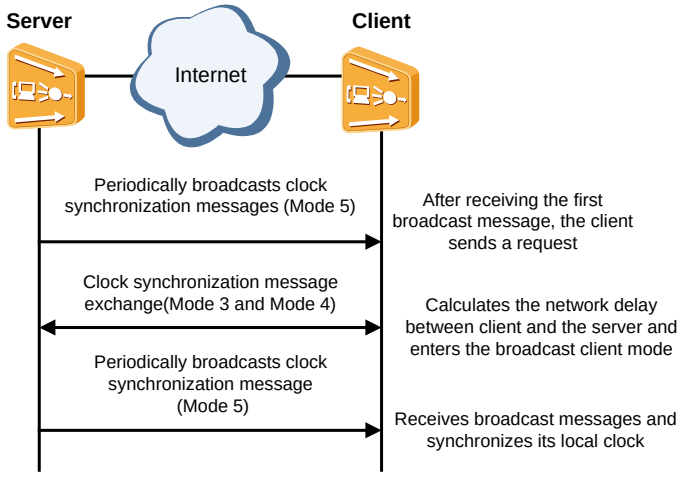


Multicast Mode

- A server running in multicast mode sends clock synchronization packets to a multicast address periodically. The value of the Mode field in a packet is set to 5. This indicates that the packet is sent by a host that runs in multicast mode. The host running in multicast mode is usually a time server running high-speed broadcast media on the network, which provides synchronization information for all of its peers but does not alter the clock of its own.
- The client listens to the multicast packets from the server. When the client receives the first broadcast packet, the client and the server exchange NTP packets whose values of Mode fields are 3 (sent by the client) and the NTP packets whose values of Mode fields are 4 (sent by the server). In this process, the client enables the server/client mode for a short time to exchange information with the remote server. This allows the client to obtain the network delay between the client and the server. Then, the client returns the multicast mode, and continues to sense the incoming multicast packets to synchronize the local clock.

Multicast mode is useful when there are large numbers of clients distributed in a network. This normally results in large number of NTP packets in the network. In the multicast mode, a single NTP multicast packet can potentially reach all the clients in the network and thus reduce the control traffic on the network.

Figure 32-45 Multicast mode

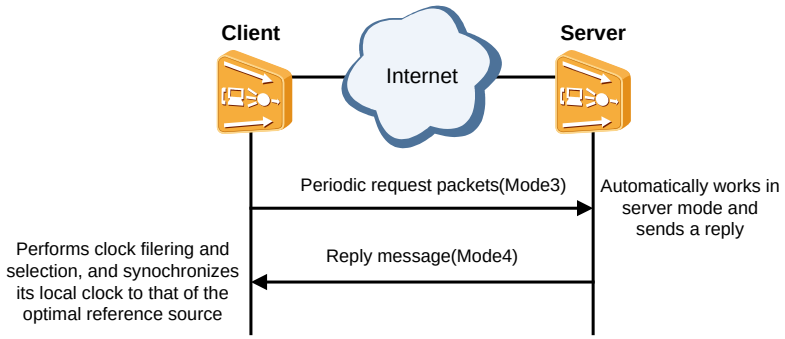


Manycast Mode

- A client operating in manycast mode, sends periodic request packets to a designated IPv4 or IPv6 multicast address in order to search for NTP server. It starts with a time-to-live (TTL) value equal to one and continuously adding one to it. A designated manycast server listens for packets with that address. If a server receives the packets from the client, it returns an ordinary server (mode 4) packet to the client. If client receives the packets from the server, temporary C/S connection between them is established. After a certain period of time, the client stops the search process and select the best connection from all connections. The other connection not be selected will be aging out.

Manycast mode is applied to a small set of servers scattered over the network. Clients can discover and synchronize to the closest manycast server. Manycast can especially be used where the identity of the server is not fixed and a change of server does not require reconfiguration of all the clients in the network.

Figure 32-46 Manycast mode



Kiss-o'-Death (KOD)

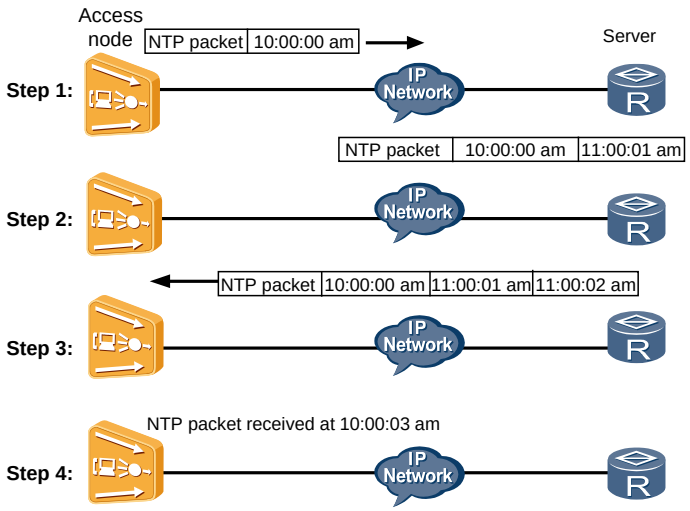
KOD packets provide useful information to a client and are used for status reporting and access control. When KOD is enabled at the server, the server may send packets with kiss codes DENY and RATE to the client.

- When the client receives packet with kiss code DENY, the client demobilizes any associations with that server and stops sending packets to that server.
- When the client receives packet with kiss code RATE, the client immediately reduces its polling interval to that server and continues to reduce it each time it receives a RATE kiss code.

NTP Working Principle

As shown in Figure 32-47, the MA5600T/MA5603T/MA5608T serves as the NTP client and the router serves as the NTP server. The MA5600T/MA5603T/MA5608T uses the time of the router as the reference and synchronizes its time with the router through NTP.

Figure 32-47 Operating principle of NTP



1. The MA5600T/MA5603T/MA5608T sends an NTP packet to the router. This packet contains the timestamp when it leaves the MA5600T/MA5603T/MA5608T. Assume that the timestamp is 10:00:00 am (T1).
2. When the NTP packet arrives at the router, the router adds its timestamp to the packet. Assume that the timestamp is 11:00:01 am (T2).
3. When the NTP packet leaves the router, the router adds another timestamp to the packet. Assume that the timestamp is 11:00:02 am (T3).
4. When the MA5600T/MA5603T/MA5608T receives the response packet, it adds a new timestamp to the packet. Assume that the timestamp is 10:00:03 am (T4).

Now, the MA5600T/MA5603T/MA5608T has sufficient information to calculate two important parameters:

- The delay for a round trip of the NTP packet=(T4-T1)-(T3-T2).
- Offset between the MA5600T/MA5603T/MA5608T and the router=((T2-T1)-(T4-T3))/2

In this way, the MA5600T/MA5603T/MA5608T can set its clock according to the information and thus keeps its clock synchronized with that of the router.

NTP Security Mechanism

When a time server in the subnet is faulty or data is maliciously modified or destroyed, timekeeping on other time servers in the subnet should not be affected. To meet this requirement, NTP provides two security mechanisms: access right and NTP authentication to guarantee the network security.

Access Right Control

You can run the **ntp-service access** command to set the right to access the NTP service of the MA5600T/MA5603T/MA5608T so as to protect the NTP service.

The MA5600T/MA5603T/MA5608T supports five levels of access rights, as shown in the following table.

Table 32-14 NTP access rights supported by the MA5600T/MA5603T/MA5608T.

Access Right	Description
peer	Indicates the minimum access right. The remote end can perform time requests and control queries for the local NTP service. The local clock can also be synchronized with the clock of the remote server.
server	Indicates that the remote end can perform time requests and control queries for the local NTP service. The local clock, however, cannot be synchronized with the clock of the remote server.
synchronization	Indicates that the remote end can perform time requests only for the local NTP service.
query	Indicates the maximum access right. The remote end can perform control queries only for the local NTP service.
limited	Controls the incoming packet rate and kiss code is sent when KoD is enabled.

NTP Authentication

The NTP authentication can be used on security-critical networks. If the NTP authentication is enabled, the MA5600T/MA5603T/MA5608T uses MD5-algorithm or HMAC-SHA256-algorithm keys to authenticate users who access the NTP service. You are recommended to use the HMAC-SHA256 algorithm to improve security. The NTP authentication must be configured on both the server and the client.

The key on the client must match the one on the server for the user to pass the NTP authentication.

The process for configuring the NTP authentication on the server and the client is as follows:

1. Enable the NTP authentication.
2. Configure the NTP authentication key.
3. Declare that the NTP authentication key is trustworthy.

NOTE

- The client is synchronized to only the server that provides the reliable key. If the key provided by the server is unreliable, the client is not synchronized to the server.
- When the client is configured with the NTP authentication, if only the server is configured with same authentication key as the client, the client can pass the authentication. Here, the server does not need to enable the NTP authentication function or declare that the key is reliable.

32.8.3 Configuring the Network Time

Configure the NTP protocol to keep the time of all devices in the network synchronized, so that the MA5600T/MA5603T/MA5608T implements various service applications based on universal time, such as the network management system and the network accounting system.

Context

Introduction to the NTP Protocol:

- The Network Time Protocol (NTP) is an application layer protocol defined in RFC 1305, which is used to synchronize the times of the distributed time server and the client. The RFC defines the structures, arithmetics, entities and protocols used in the implementation of NTP.
- NTP is developed from the time protocol and the ICMP timestamp message protocol, with special design on the aspects of accuracy and robustness.
- NTP runs over UDP with port number as 123.
- Any local system that runs NTP can be time synchronized by other clock sources, and also act as a clock source to synchronize other clocks. In addition, mutual synchronization can be done through NTP packets exchanges.

NTP is applied to the following situations where all the clocks of hosts or routers in a network need to be consistent:

- In the network management, an analysis of log or debugging information collected from different routers needs time for reference.
- The charging system requires the clocks of all devices to be consistent.
- Completing certain functions, for example, timing restart of all the routers in a network requires the clocks of all the routers be consistent.
- When several systems work together on the same complicate event, they have to take the same clock for reference to ensure a correct implementation order.
- Incremental backup between the backup server and clients requires clocks on them be synchronized.

When all the devices on a network need to be synchronized, it is almost impossible for an administrator to manually change the system clock by using a command line. This is because the work load is heavy and clock accuracy cannot be ensured. NTP can quickly synchronize the clocks of network devices and ensure their precision.

There are four NTP modes: server/client, peer, broadcast and multicast modes. The MA5600T/MA5603T/MA5608T supports all these modes.

Default Configuration

Table 32-15 provides the default configuration for NTP.

Table 32-15 Default configuration for NTP

Parameter	Default Value
NTP-service authentication function	Disable
NTP-service authentication key	None
The maximum allowed number of sessions	100
Clock stratum	16

(Optional) Configuring NTP Authentication

This topic describes how to configure NTP authentication to improve the network security and prevent unauthorized users from modifying the clock.

Prerequisites

Before configuring the NTP client/server mode, make sure that the network interface and the routing protocol of the MA5600T/MA5603T/MA5608T are configured so that the server and the client are reachable to each other at the network layer.

Context

In certain networks that have strict requirements on security, enable NTP authentication when running the NTP protocol. Configuring NTP authentication is classified into configuring NTP authentication on the client and configuring NTP authentication on the server.

Precaution

- If NTP authentication is not enabled on the client, the client can synchronize with the server, regardless of whether NTP authentication is enabled on the server.
- If NTP authentication is enabled, a reliable key should be configured.
- The configuration of the server must be the same as that of the client.
- When NTP authentication is enabled on the client, the client can pass the authentication if the server is configured with the same key as that of the client. In this case, you do not need to enable NTP authentication on the server or declare that the key is reliable.
- The client synchronizes with only the server that provides the reliable key. If the key provided by the server is unreliable, the client does not synchronize with the server.
- The flow of configuring NTP authentication is as follows: start->enable NTP authentication->configure the reliable NTP authentication key->declare the reliable key->end.

Procedure

Run the **undo ntp-service server disable** command to enable NTP server functionality.

Step 1 Run the **ntp-service authentication enable** command to enable NTP authentication.

Step 2 Run the **ntp-service authentication-keyid** command to set an NTP authentication key.

Step 3 Run the **ntp-service reliable authentication-keyid** command to declare that the key is reliable.

----End

Example

To enable NTP authentication, set the NTP authentication key as **aNiceKey** with the key number 42, and then define key 42 as a reliable key, do as follows:

```
huawei(config)#undo ntp-service server disable
huawei(config)#ntp-service authentication enable
huawei(config)#ntp-service authentication-keyid 42 authentication-mode md5 aNice
Key
huawei(config)#ntp-service reliable authentication-keyid 42
```

Configuring the NTP Broadcast Mode

This topic describes how to configure the MA5600T/MA5603T/MA5608T for clock synchronization in the NTP broadcast mode. After the configuration is completed, the server periodically broadcasts clock synchronization packets through a specified port, and the client listens to the broadcast packets sent from the server and synchronizes the local clock according to the received broadcast packets.

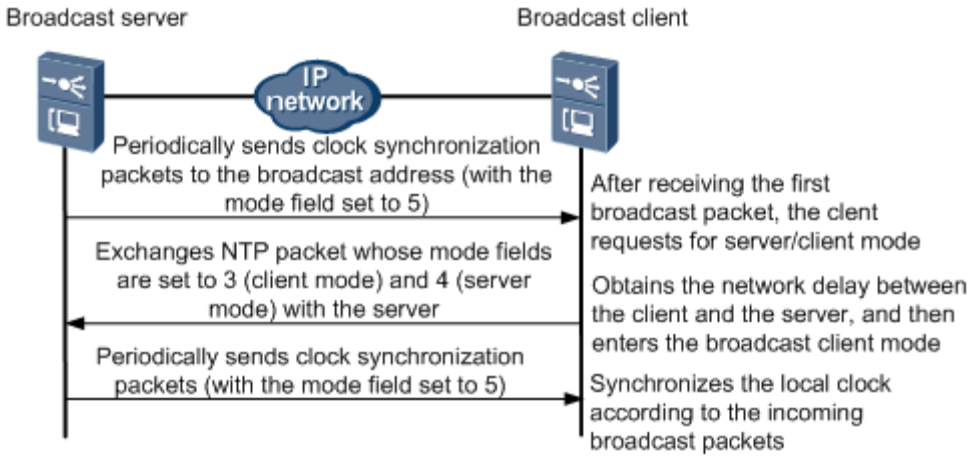
Prerequisites

Before configuring the NTP broadcast mode, make sure that the network interface and the routing protocol of the MA5600T/MA5603T/MA5608T are configured so that the server and the client are reachable to each other at the network layer.

Context

In the broadcast mode, the server periodically sends clock synchronization packets to the broadcast address 255.255.255.255 or 255:255::255:255, with the mode field set to 5 (indicating the broadcast mode). The client listens to the broadcast packets sent from the server. After receiving the first broadcast packet, the client exchanges NTP packet whose mode fields are set to 3 (client mode) and 4 (server mode) with the server to estimate the network delay between the client and the server. The client then enters the broadcast client mode, continues to listen to the incoming broadcast packets, and synchronizes the local clock according to the incoming broadcast packets, as shown in Figure 32-48.

Figure 32-48 NTP broadcast mode



Precaution

1. In the broadcast mode, you should configure both the NTP server and the NTP client.
2. The clock stratum of the synchronizing device must be higher than or equal to that of the synchronized device. Otherwise, the clock synchronization fails.

Procedure

- Configure the NTP broadcast server host.
  a. (Optional) In IPv6 systems, run the **ipv6** command to globally enable IPv6.
  b. Enable the NTP server functionality.
     - Run the **undo ntp-service server disable** command to enable IPv4 NTP server functionality.
     - Run the **undo ntp-service ipv6 server disable** command to enable IPv6 NTP server functionality.
  c. Run the **ntp-service refclock-master** command to configure the local clock as the master NTP clock, and specify the stratum of the master NTP clock.
  d. (Optional) Configure NTP authentication.

     In certain networks that have strict requirements on security, it is recommended that you enable NTP authentication when running the NTP protocol. The configuration of the server must be the same as that of the client.

     i. Run the **ntp-service authentication enable** command to enable NTP authentication.
     ii. Run the **ntp-service authentication-keyid** command to set an NTP authentication key.
     iii. Run the **ntp-service reliable authentication-keyid** command to declare that the key is reliable.
  e. Add a VLAN Layer 3 interface.
     i. Run the **vlan** command to create a VLAN.
     ii. Run the **port vlan** command to add an upstream port to the VLAN so that the user packets carrying the VLAN tag are transmitted upstream through the upstream port.

iii. In the global config mode, run the **interface vlan** command to create a VLAN interface, and then enter the VLAN interface mode to configure the Layer 3 interface.

iv. Configure the IP address based on the requirement.

- Run the **ip address** command to configure the IPv4 address and subnet mask of the VLAN interface so that the IP packets in the VLAN can participate in the Layer 3 forwarding.
- Run the **ipv6 address** command to configure the IPv6 address and its subnet prefix length so that the IP packets in the VLAN can participate in the Layer 3 forwarding.

f. Run the **ntp-service broadcast-server** command to configure the NTP broadcast server mode of the host, and specify the key ID for the server to send packets to the client.

- Configure the NTP broadcast client host.

a. (Optional) Configure NTP authentication.

In certain networks that have strict requirements on security, it is recommended that you enable NTP authentication when running the NTP protocol. The configuration of the server must be the same as that of the client.

i. Run the **ntp-service authentication enable** command to enable NTP authentication.

ii. Run the **ntp-service authentication-keyid** command to set an NTP authentication key.

iii. Run the **ntp-service reliable authentication-keyid** command to declare that the key is reliable.

b. Add a VLAN Layer 3 interface.

i. Run the **vlan** command to create a VLAN.

ii. Run the **port vlan** command to add an upstream port to the VLAN so that the user packets carrying the VLAN tag are transmitted upstream through the upstream port.

iii. In the global config mode, run the **interface vlan** command to create a VLAN interface, and then enter the VLAN interface mode to configure the Layer 3 interface.

iv. Configure the IP address based on the requirement.

- Run the **ip address** command to configure the IPv4 address and subnet mask of the VLAN interface so that the IP packets in the VLAN can participate in the Layer 3 forwarding.
- Run the **ipv6 address** command to configure the IPv6 address and its subnet prefix length so that the IP packets in the VLAN can participate in the Layer 3 forwarding.

c. Run the **ntp-service broadcast-client** command to configure a host as the NTP broadcast client.

----End

Example

(IPv4) Assume the following configurations: MA5600T/MA5603T/MA5608T_S uses the local clock as the master NTP clock on stratum 2 and works in the NTP broadcast mode, broadcasting clock synchronization packets periodically through IP address 10.10.10.10/24 of

the Layer 3 interface of VLAN 2, and MA5600T/MA5603T/MA5608T_C functions as the NTP client, listening to the broadcast packets sent from the server through IP address 10.10.10.20/24 of the Layer 3 interface of VLAN 2 and synchronizing with the clock on the broadcast server. To perform these configurations, do as follows:

1. On MA5600T/MA5603T/MA5608T_S:

```
huawei(config)#undo ntp-service server disable
huawei(config)#ntp-service refclock-master 2
huawei(config)#vlan 2 standard
huawei(config)#port vlan 2 0/19 0
huawei(config)#interface vlanif 2
huawei(config-if-vlanif2)#ip address 10.10.10.10 24
huawei(config-if-vlanif2)#ntp-service broadcast-server
huawei(config-if-vlanif2)#quit
```

2. On MA5600T/MA5603T/MA5608T_C:

```
huawei(config)#undo ntp-service server disable
huawei(config)#vlan 2 standard
huawei(config)#port vlan 2 0/19 0
huawei(config)#interface vlanif 2
huawei(config-if-vlanif2)#ip address 10.10.10.20 24
huawei(config-if-vlanif2)#ntp-service broadcast-client
huawei(config-if-vlanif2)#quit
```

(IPv6) Assume the following configurations: MA5600T/MA5603T/MA5608T_S uses the local clock as the master NTP clock on stratum 2 and works in the NTP broadcast mode, broadcasting clock synchronization packets periodically through IP address 10:10::10:10/24 of the Layer 3 interface of VLAN 2, and MA5600T/MA5603T/MA5608T_C functions as the NTP client, listening to the broadcast packets sent from the server through IP address 10.10.10.20/24 of the Layer 3 interface of VLAN 2 and synchronizing with the clock on the broadcast server. To perform these configurations, do as follows:

1. On MA5600T/MA5603T/MA5608T_S:

```
huawei(config)#ipv6
huawei(config)#undo ntp-service ipv6 server disable
huawei(config)#ntp-service refclock-master 2
huawei(config)#vlan 2 standard
huawei(config)#port vlan 2 0/19 0
huawei(config)#interface vlanif 2
huawei(config-if-vlanif2)#ipv6 enable
huawei(config-if-vlanif2)#ipv6 address 10:10::10:10 24
huawei(config-if-vlanif2)#ntp-service broadcast-server
huawei(config-if-vlanif2)#quit
```

2. On MA5600T/MA5603T/MA5608T_C:

```
huawei(config)#ipv6
huawei(config)#undo ntp-service ipv6 server disable
huawei(config)#vlan 2 standard
huawei(config)#port vlan 2 0/19 0
huawei(config)#interface vlanif 2
huawei(config-if-vlanif2)#ipv6 enable
huawei(config-if-vlanif2)#ipv6 address 10:10::10:20 24
huawei(config-if-vlanif2)#ntp-service broadcast-client
```

```
huawei(config-if-vlanif2)#quit
```

Configuring the NTP Multicast Mode

This topic describes how to configure the MA5600T/MA5603T/MA5608T for clock synchronization in the NTP multicast mode. After the configuration is completed, the server periodically multicasts clock synchronization packets through a specified port, and the client listens to the multicast packets sent from the server and synchronizes the local clock according to the received multicast packets.

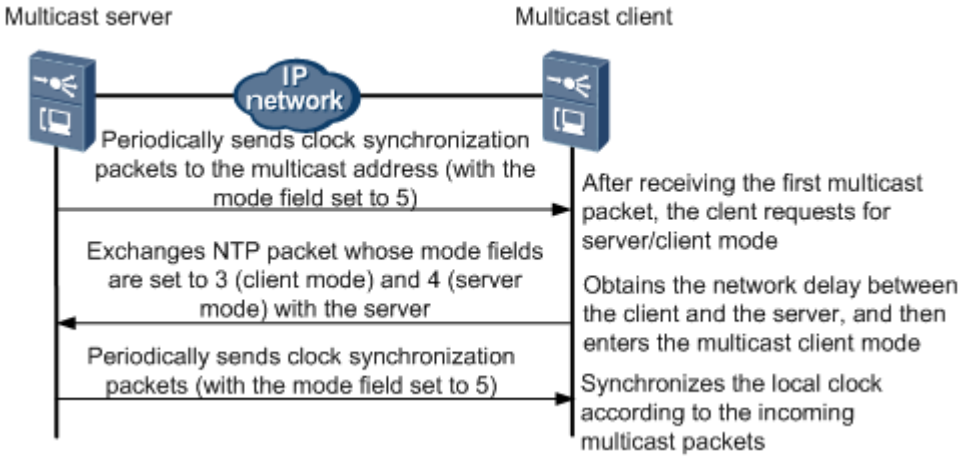
Prerequisites

Before configuring the NTP multicast mode, make sure that the network interface and the routing protocol of the MA5600T/MA5603T/MA5608T are configured so that the server and the client are reachable to each other at the network layer.

Context

In the multicast mode, the server periodically sends clock synchronization packets to the multicast address configured by the user. The default NTP multicast address 224.0.1.1 is used if the multicast address is not configured. The mode field of clock synchronization packet is set to 5 (multicast mode). The client listens to the multicast packets sent from the server. After receiving the first multicast packet, the client exchanges NTP packet whose mode fields are set to 3 (client mode) and 4 (server mode) with the server to estimate the network delay between the client and the server. The client then enters the multicast client mode, continues to listen to the incoming multicast packets, and synchronizes the local clock according to the incoming multicast packets, as shown in Figure 32-49.

Figure 32-49 NTP multicast mode



Precaution

1. In the multicast mode, you should configure both the NTP server and the NTP client.
2. The clock stratum of the synchronizing device must be higher than or equal to that of the synchronized device. Otherwise, the clock synchronization fails.

Procedure

- Configure the NTP multicast server host.
  a. (Optional) In IPv6 systems, run the **ipv6** command to globally enable IPv6.
  b. Run the **ntp-service refclock-master** command to configure the local clock as the master NTP clock, and specify the stratum of the master NTP clock.
  c. (Optional) Configure NTP authentication.

     In certain networks that have strict requirements on security, it is recommended that you enable NTP authentication when running the NTP protocol. The configuration of the server must be the same as that of the client.

     i. Run the **ntp-service authentication enable** command to enable NTP authentication.
     ii. Run the **ntp-service authentication-keyid** command to set an NTP authentication key.
     iii. Run the **ntp-service reliable authentication-keyid** command to declare that the key is reliable.
  d. Add a VLAN Layer 3 interface.
     i. Run the **vlan** command to create a VLAN.
     ii. Run the **port vlan** command to add an upstream port to the VLAN so that the user packets carrying the VLAN tag are transmitted upstream through the upstream port.
     iii. In the global config mode, run the **interface vlan** command to create a VLAN interface, and then enter the VLAN interface mode to configure the Layer 3 interface.
     iv. Configure the IP address based on the requirement.
        - Run the **ip address** command to configure the IPv4 address and subnet mask of the VLAN interface so that the IP packets in the VLAN can participate in the Layer 3 forwarding.
        - Run the **ipv6 address** command to configure the IPv6 address and its subnet prefix length so that the IP packets in the VLAN can participate in the Layer 3 forwarding.
  e. Run the **ntp-service multicast-server** command to configure the NTP multicast server mode of the host, and specify the key ID for the server to send packets to the client.
- Configure the NTP multicast client host.
  a. (Optional) Configure NTP authentication.

     In certain networks that have strict requirements on security, it is recommended that you enable NTP authentication when running the NTP protocol. The configuration of the server must be the same as that of the client.

     i. Run the **ntp-service authentication enable** command to enable NTP authentication.
     ii. Run the **ntp-service authentication-keyid** command to set an NTP authentication key.
     iii. Run the **ntp-service reliable authentication-keyid** command to declare that the key is reliable.
  b. Add a VLAN Layer 3 interface.
     i. Run the **vlan** command to create a VLAN.

ii. Run the **port vlan** command to add an upstream port to the VLAN so that the user packets carrying the VLAN tag are transmitted upstream through the upstream port.

iii. In the global config mode, run the **interface vlan** command to create a VLAN interface, and then enter the VLAN interface mode to configure the Layer 3 interface.

iv. Configure the IP address based on the requirement.

- Run the **ip address** command to configure the IPv4 address and subnet mask of the VLAN interface so that the IP packets in the VLAN can participate in the Layer 3 forwarding.
- Run the **ipv6 address** command to configure the IPv6 address and its subnet prefix length so that the IP packets in the VLAN can participate in the Layer 3 forwarding.

c. Run the **ntp-service multicast-client** command to configure a host as the NTP multicast client.

----End

Example

(IPv4) Assume the following configurations: MA5600T/MA5603T/MA5608T_S uses the local clock as the master NTP clock on stratum 2 and works in the NTP multicast mode, multicasting clock synchronization packets periodically through IP address 10.10.10.10/24 of the Layer 3 interface of VLAN 2, and MA5600T/MA5603T/MA5608T_C functions as the NTP client, listening to the multicast packets sent from the server through IP address 10.10.10.20/24 of the Layer 3 interface of VLAN 2 and synchronizing with the clock on the multicast server. To perform these configurations, do as follows:

1. On MA5600T/MA5603T/MA5608T_S:

```
huawei(config)#ntp-service refclock-master 2
huawei(config)#vlan 2 standard
huawei(config)#port vlan 2 0/19 0
huawei(config)#interface vlanif 2
huawei(config-if-vlanif2)#ip address 10.10.10.10 24
huawei(config-if-vlanif2)#ntp-service multicast-server
huawei(config-if-vlanif2)#quit
```

2. On MA5600T/MA5603T/MA5608T_C:

```
huawei(config)#vlan 2 standard
huawei(config)#port vlan 2 0/19 0
huawei(config)#interface vlanif 2
huawei(config-if-vlanif2)#ip address 10.10.10.20 24
huawei(config-if-vlanif2)#ntp-service multicast-client
huawei(config-if-vlanif2)#quit
```

(IPv6) Assume the following configurations: MA5600T/MA5603T/MA5608T_S uses the local clock as the master NTP clock on stratum 2 and works in the NTP multicast mode, multicasting clock synchronization packets periodically through IP address 10:10::10:10/24 of the Layer 3 interface of VLAN 2, and MA5600T/MA5603T/MA5608T_C functions as the NTP client, listening to the multicast packets sent from the server through IP address 10:10::10:20/24 of the Layer 3 interface of VLAN 2 and synchronizing with the clock on the multicast server. To perform these configurations, do as follows:

1. On MA5600T/MA5603T/MA5608T_S:

```
huawei(config)#ntp-service refclock-master 2
huawei(config)#vlan 2 standard
huawei(config)#port vlan 2 0/19 0
huawei(config)#interface vlanif 2
huawei(config-if-vlanif2)#ipv6 address 10:10::10:10 24
huawei(config-if-vlanif2)#ntp-service multicast-server
huawei(config-if-vlanif2)#quit
```

2. On MA5600T/MA5603T/MA5608T_C:

```
huawei(config)#vlan 2 standard
huawei(config)#port vlan 2 0/19 0
huawei(config)#interface vlanif 2
huawei(config-if-vlanif2)#ipv6 address 10:10::10:20 24
huawei(config-if-vlanif2)#ntp-service multicast-client
huawei(config-if-vlanif2)#quit
```

Configuring the Unicast NTP Client

This topic describes how to configure the MA5600T/MA5603T/MA5608T as the NTP client to synchronize with the NTP server in the network.

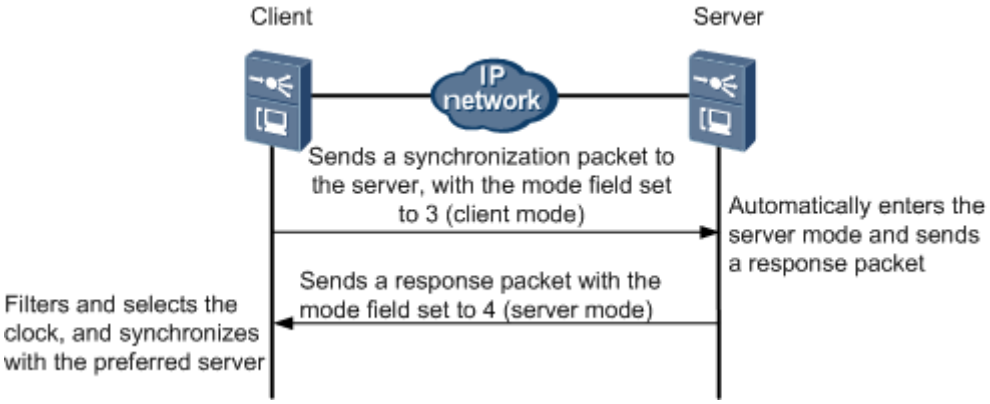
Prerequisites

Before configuring the NTP client/server mode, make sure that the network interface and the routing protocol of the MA5600T/MA5603T/MA5608T are configured so that the server and the client are reachable to each other at the network layer.

Context

In the client/server mode, the client sends a synchronization packet to the server, with the mode field set to 3 (client mode). After receiving the packet, the server automatically enters the server mode and sends a response packet with the mode field set to 4 (server mode). After receiving the response from the server, the client filters and selects the clock, and synchronizes with the preferred server, as shown in Figure 32-50.

Figure 32-50 NTP client/server mode



Precaution

1. In the client/server mode, you need to configure only the client, and do not need to configure the server.
2. The clock stratum of the synchronizing device must be lower than or equal to that of the synchronized device. Otherwise, the clock synchronization fails.

Procedure

Add a VLAN Layer 3 interface.

1. (Optional) In IPv6 systems, run the **ipv6** command to globally enable IPv6.
2. Run the **vlan** command to create a VLAN.
3. Run the **port vlan** command to add an upstream port to the VLAN so that the user packets carrying the VLAN tag are transmitted upstream through the upstream port.
4. In the global config mode, run the **interface vlan** command to create a VLAN interface, and then enter the VLAN interface mode to configure the Layer 3 interface.
5. Configure the IP address based on the system.
   - Run the **ip address** command to configure the IPv4 address and subnet mask of the VLAN interface so that the IP packets in the VLAN can participate in the Layer 3 forwarding.
   - Run the **ipv6 address** command to configure the IPv6 address and its subnet prefix length so that the IP packets in the VLAN can participate in the Layer 3 forwarding.

Step 1 Enable the NTP server functionality.

- Run the **undo ntp-service server disable** command to enable IPv4 NTP server functionality.
- Run the **undo ntp-service ipv6 server disable** command to enable IPv6 NTP server functionality.

Step 2 Run the **ntp-service unicast-server** command to configure the NTP unicast server mode, and specify the IP address of the remote server that functions as the local timer server and the interface for transmitting and receiving NTP packets.

NOTE

- In this command, *ip-address* is a unicast address, which cannot be a broadcast address, a multicast address, or the IP address of a local clock.
- After the source interface of the NTP packets is specified by *source-interface*, the source IP address of the NTP packets is configured as the primary IP address of the specified interface.
- A server can function as a time server to synchronize other devices only after its clock is synchronized.
- When the clock stratum of the server is higher than or equal to that of the client, the client does not synchronize with the server.

You can run the **ntp-service unicast-server** command for multiple times to configure multiple servers. Then, the client selects the best server according to clock priorities or preemption. For example, assuming

- There are A and B two servers.
- A sets up a temporary NTP session, then B sets up another.

When the condition (includes the level and the precision) of A and B are the same, Table 32-16 shows the clock source.

Table 32-16 The clock source option

Mode (A)	Mode (B)	Clock Source
None	None	A
None	Priority	B
None	Preemption	A
Priority	None	A
Priority	Priority	A
Priority	Preemption	A
Preemption	None	B
Preemption	Priority	B
Preemption	Preemption	A

Step 3 (Optional) Configure the ACL rules.

Filter the packets that pass through the Layer 3 interface. Only the IP packet from the clock server is allowed to access the Layer 3 interface. Other unauthorized packets are not allowed to access the Layer 3 interface. It is recommended to use the ACL rules for the system that has high requirements on security.

1. Create the ACL based on the Requirement.
   - Run the **acl** *adv-acl-numbe* command to create an ACL for the IPv4 system.
   - Run the **acl ipv6** *adv-acl-numbe* command to create an ACL for the IPv6 system.
2. Run the **rule** command to classify traffic according to the source IP address, destination IP address, type of the protocol over IP, and features or protocol of the packet, allowing or forbidding the data packets that meet related conditions to pass.
3. Run the **packet-filter** command to configure an ACL filtering rule for a specified port, and make the configuration take effect.

----End

Example

(IPv4) Assume the following configurations: One MA5600T/MA5603T/MA5608T functions as the NTP server (IP address: 10.20.20.20/24), the other MA5600T/MA5603T/MA5608T (IP address of the Layer 3 interface of VLAN 2: 10.10.10.10/24, gateway IP address: 10.10.10.1) functions as the NTP client, the NTP client sends the clock synchronization request packet through the VLAN Layer 3 interface to the NTP server, the NTP server responds to the request packet, and ACL rules are configured to allow only IP packets from the clock server to access the Layer 3 interface. To perform these configurations, do as follows:

```
huawei(config)#vlan 2 standard
huawei(config)#port vlan 2 0/19 0
huawei(config)#interface vlanif 2
huawei(config-if-vlanif2)#ip address 10.10.10.10 24
huawei(config-if-vlanif2)#quit
huawei(config)#undo ntp-service server disable
```

```
huawei(config)#ntp-service unicast-server 10.20.20.20 source-interface vlanif 2
huawei(config)#acl 3010
huawei(config-acl-adv-3010)#rule deny ip source any destination 10.10.10.10
 0.0.0.0
huawei(config-acl-adv-3010)#rule permit ip source 10.20.20.20 0.0.0.0 destination
 10.10.10.10 0.0.0.0
huawei(config-acl-adv-3010)#quit
huawei(config)#packet-filter inbound ip-group 3010 port 0/19/0
```

(IPv6) Assume the following configurations: One MA5600T/MA5603T/MA5608T functions as the NTP server (IP address: 10:20::20:20/24), the other MA5600T/MA5603T/MA5608T (IP address of the Layer 3 interface of VLAN 2: 10:10::10:10/24, gateway IP address: 10:10::10:1) functions as the NTP client, the NTP client sends the clock synchronization request packet through the VLAN Layer 3 interface to the NTP server, the NTP server responds to the request packet, and ACL rules are configured to allow only IP packets from the clock server to access the Layer 3 interface. To perform these configurations, do as follows:

```
huawei(config)#ipv6
huawei(config)#vlan 2 standard
huawei(config)#port vlan 2 0/19 0
huawei(config)#interface vlanif 2
huawei(config-if-vlanif2)#ipv6 enable
huawei(config-if-vlanif2)#ipv6 address 10:10::10:10 24
huawei(config-if-vlanif2)#quit
huawei(config)#undo ntp-service ipv6 server disable
huawei(config)#ntp-service unicast-server ipv6 10:20::20:20 source-interface vlanif
2
huawei(config)#acl ipv6 3010
huawei(config-acl6-adv-3010)#rule deny ip source any destination 10:10::10:10
 0:0::0:0
huawei(config-acl6-adv-3010)#rule permit ip source 10:20::20:20 0:0::0:0 destination
 10:10::10:10 0:0::0:0
huawei(config-acl6-adv-3010)#quit
huawei(config)#packet-filter inbound ip-group 3010 port 0/19/0
```

Configuring the NTP Peer

This topic describes how to configure the MA5600T/MA5603T/MA5608T for clock synchronization in the NTP peer mode. In the peer mode, configure only the active peer, and the passive peer does not need to be configured. In the peer mode, the active peer and the passive peer can synchronize with each other. The peer with a higher clock stratum is synchronized by the peer with a lower clock stratum.

Prerequisites

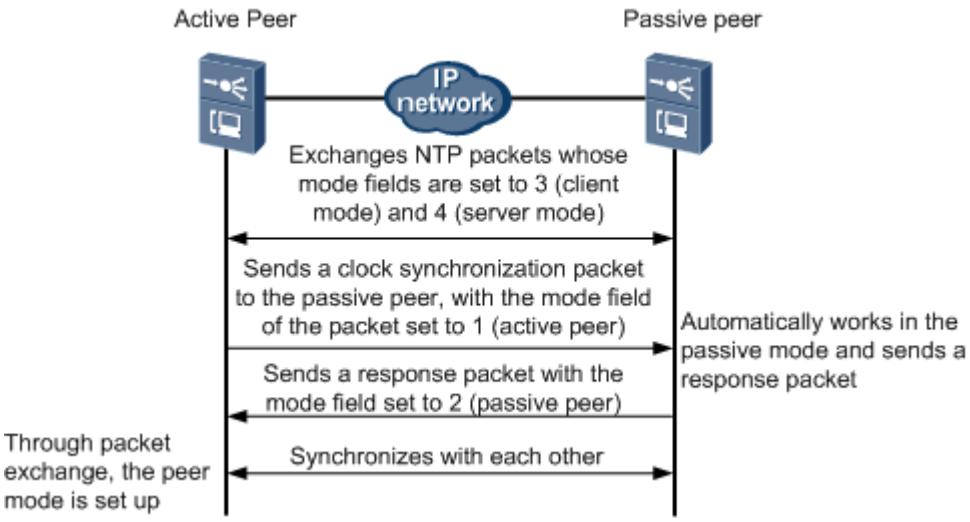
Before configuring the NTP peer mode, make sure that the network interface and the routing protocol of the MA5600T/MA5603T/MA5608T are configured so that the server and the client are reachable to each other at the network layer.

Context

In the peer mode, the active peer and the passive peer exchange NTP packets whose mode fields are set to 3 (client mode) and 4 (server mode). Then, the active peer sends a clock synchronization packet to the passive peer, with the mode field of the packet set to 1 (active

peer). After receiving the packet, the passive peer automatically works in the passive mode and sends a response packet with the mode field set to 2 (passive peer). Through packet exchange, the peer mode is set up. The active peer and the passive peer can synchronize with each other. If both the clock of the active peer and that of the passive peer are synchronized, the clock on a lower stratum is used, as shown in Figure 32-51.

Figure 32-51 NTP peer mode



Precaution

1. In the peer mode, you need to configure the NTP mode only on the active peer.
2. The peers determine clock synchronization according to the clock stratum instead of according to whether the peer is an active peer.

Procedure

Configure the NTP active peer.

1. (Optional) In IPv6 systems, run the **ipv6** command to globally enable IPv6.
2. Enable the NTP server functionality.
   - Run the **undo ntp-service server disable** command to enable IPv4 NTP server functionality.
   - Run the **undo ntp-service ipv6 server disable** command to enable IPv6 NTP server functionality.
3. Run the **ntp-service refclock-master** command to configure the local clock as the master NTP clock, and specify the stratum of the master NTP clock.
4. Run the **ntp-service unicast-peer** command to configure the NTP peer mode, and specify the IP address of the remote server that functions as the local timer server and the interface for transmitting and receiving NTP packets.

NOTE

- In this command, *ip-address* is a unicast address, which cannot be a broadcast address, a multicast address, or the IP address of a reference clock.
- After the source interface of the NTP packets is specified by *source-interface*, the source IP address of the NTP packets is configured as the primary IP address of the specified interface.

Step 1 Add a VLAN Layer 3 interface.

1. Run the **vlan** command to create a VLAN.
2. Run the **port vlan** command to add an upstream port to the VLAN so that the user packets carrying the VLAN tag are transmitted upstream through the upstream port.
3. In the global config mode, run the **interface vlan** command to create a VLAN interface, and then enter the VLAN interface mode to configure the Layer 3 interface.
4. Configure the IP address based on the system.
   - Run the **ip address** command to configure the IPv4 address and subnet mask of the VLAN interface so that the IP packets in the VLAN can participate in the Layer 3 forwarding.
   - Run the **ipv6 address** command to configure the IPv6 address and its subnet prefix length so that the IP packets in the VLAN can participate in the Layer 3 forwarding.

----End

Example

Assume the following configurations: One MA5600T/MA5603T/MA5608T functions as the NTP active peer (IP address of the Layer 3 interface of VLAN 2: 10.10.10.10/24) and works on clock stratum 4, the other MA5600T/MA5603T/MA5608T (IP address: 10.10.10.20/24) functions as the NTP passive peer, the active peer sends a clock synchronization request packet through the VLAN Layer 3 interface to the passive peer, the passive peer responds to the request packet, and the peer with a higher clock stratum is synchronized by the peer with a lower clock stratum. To perform these configurations, do as follows:

```
huawei(config)#undo ntp-service server disable
huawei(config)#ntp-service refclock-master 4
huawei(config)#ntp-service unicast-peer
huawei(config)#vlan 2 standard
huawei(config)#port vlan 2 0/19 0
huawei(config)#interface vlanif 2
huawei(config-if-vlanif2)#ip address 10.10.10.10 24
huawei(config-if-vlanif2)#quit
```

Assume the following configurations: One MA5600T/MA5603T/MA5608T functions as the NTP active peer (IP address of the Layer 3 interface of VLAN 2: 10:10::10:10/24) and works on clock stratum 4, the other MA5600T/MA5603T/MA5608T (IP address: 10:10::10:20/24) functions as the NTP passive peer, the active peer sends a clock synchronization request packet through the VLAN Layer 3 interface to the passive peer, the passive peer responds to the request packet, and the peer with a higher clock stratum is synchronized by the peer with a lower clock stratum. To perform these configurations, do as follows:

```
huawei(config)#ipv6
huawei(config)#undo ntp-service ipv6 server disable
huawei(config)#ntp-service refclock-master 4
huawei(config)#ntp-service unicast-peer ipv6 10:10::10:20 source-interface vlanif 2
huawei(config)#vlan 2 standard
huawei(config)#port vlan 2 0/19 0
huawei(config)#interface vlanif 2
huawei(config-if-vlanif2)#ipv6 enable
huawei(config-if-vlanif2)#ipv6 address 10:10::10:10 24
huawei(config-if-vlanif2)#quit
```

Configuring the NTP Manycast Mode

This topic describes how to configure the MA5600T/MA5603T/MA5608T for clock synchronization in the NTP manycast mode. The manycast mode is a mechanism used by NTP clients to dynamically discover NTP servers. After the manycast mode is enabled, unicast NTP servers do not need to be manually configured for clients.

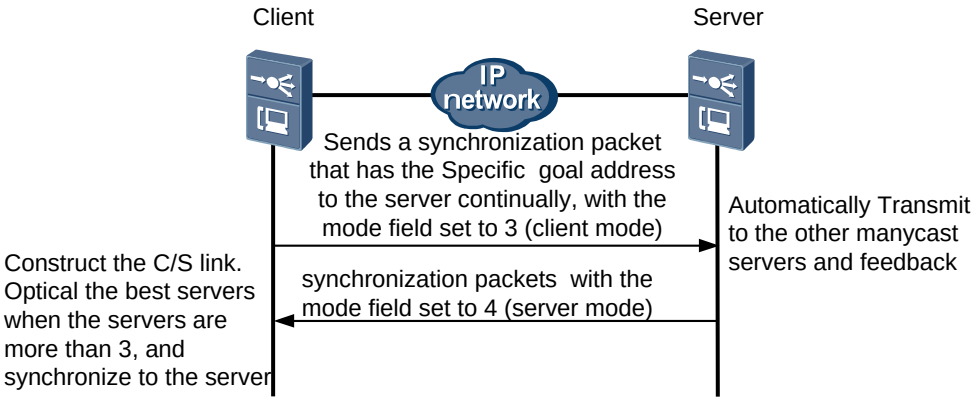
Prerequisites

Before configuring the NTP manycast mode, make sure that the network interface and the routing protocol of the MA5600T/MA5603T/MA5608T are configured so that the server and the client are reachable to each other at the network layer.

Context

In the manycast mode, manycast clients search manycast NTP servers by sending NTP packets using the specified manycast IP address as the destination IP address. The mode field in the packets is set to 3 (client mode). manycast NTP servers listen to the manycast IP address, and send response packets after receiving the packets. The mode field in the response packets is set to 4 (server mode). After receiving the response packets, the manycast clients set up temporary client/server (C/S) connections. After setting up a certain number of C/S connections (the required number of C/S connections have a fixed value of 3), the manycast clients stop searching and select the optimal NTP server from the C/S connections. The unselected C/S connections will be aged. Figure 32-52 shows the NTP manycast mode.

Figure 32-52 NTP manycast mode



Precaution

1. In the manycast mode, you should configure both the NTP server and the NTP client.
2. The clock stratum of the synchronizing device must be higher than or equal to that of the synchronized device. Otherwise, the clock synchronization fails.

Procedure

- Configure the NTP manycast server host.
  a. (Optional) In IPv6 systems, run the **ipv6** command to globally enable IPv6.
  b. Enable the NTP server functionality.

- Run the **undo ntp-service server disable** command to enable IPv4 NTP server functionality.
- Run the **undo ntp-service ipv6 server disable** command to enable IPv6 NTP server functionality.

c. Run the **ntp-service refclock-master** command to configure the local clock as the master NTP clock, and specify the stratum of the master NTP clock.

d. (Optional) Configure NTP authentication.

In certain networks that have strict requirements on security, it is recommended that you enable NTP authentication when running the NTP protocol. The configuration of the server must be the same as that of the client.

i. Run the **ntp-service authentication enable** command to enable NTP authentication.

ii. Run the **ntp-service authentication-keyid** command to set an NTP authentication key.

iii. Run the **ntp-service reliable authentication-keyid** command to declare that the key is reliable.

e. Add a VLAN Layer 3 interface.

i. Run the **vlan** command to create a VLAN.

ii. Run the **port vlan** command to add an upstream port to the VLAN so that the user packets carrying the VLAN tag are transmitted upstream through the upstream port.

iii. In the global config mode, run the **interface vlan** command to create a VLAN interface, and then enter the VLAN interface mode to configure the Layer 3 interface.

iv. Configure the IP address based on the requirement.

- Run the **ip address** command to configure the IPv4 address and subnet mask of the VLAN interface so that the IP packets in the VLAN can participate in the Layer 3 forwarding.
- Run the **ipv6 address** command to configure the IPv6 address and its subnet prefix length so that the IP packets in the VLAN can participate in the Layer 3 forwarding.

f. Run the **ntp-service manycast-server** command to configure the NTP manycast server mode of the host, and specify the key ID for the server to send packets to the client.

- Configure the NTP manycast client host.

a. (Optional) Configure NTP authentication.

In certain networks that have strict requirements on security, it is recommended that you enable NTP authentication when running the NTP protocol. The configuration of the server must be the same as that of the client.

i. Run the **ntp-service authentication enable** command to enable NTP authentication.

ii. Run the **ntp-service authentication-keyid** command to set an NTP authentication key.

iii. Run the **ntp-service reliable authentication-keyid** command to declare that the key is reliable.

b. Add a VLAN Layer 3 interface.

i. Run the **vlan** command to create a VLAN.

ii. Run the **port vlan** command to add an upstream port to the VLAN so that the user packets carrying the VLAN tag are transmitted upstream through the upstream port.

iii. In the global config mode, run the **interface vlan** command to create a VLAN interface, and then enter the VLAN interface mode to configure the Layer 3 interface.

iv. Configure the IP address based on the requirement.

- Run the **ip address** command to configure the IPv4 address and subnet mask of the VLAN interface so that the IP packets in the VLAN can participate in the Layer 3 forwarding.
- Run the **ipv6 address** command to configure the IPv6 address and its subnet prefix length so that the IP packets in the VLAN can participate in the Layer 3 forwarding.

c. Run the **ntp-service manycast-client** command to configure a host as the NTP manycast client.

----End

Example

(IPv4) Assume the following configurations: MA5600T/MA5603T/MA5608T_S uses the local clock as the master NTP clock on stratum 2 and works in the NTP manycast mode, listening to clock synchronization packets that use 10.10.10.10/24 as the destination IP address through IP address 10.10.10.10/24 of the L3 interface of VLAN 2 and establishing C/S connections with manycast clients, and MA5600T/MA5603T/MA5608T_C functions as the NTP manycast client, manycasting clock synchronization packets that use 10.10.10.10/24 as the destination IP address continousely through IP address 10.10.10.10/24 of the L3 interface of VLAN 2 and synchronizing with the clock on the manycast server by establishing C/S connections with the manycast server. To perform these configurations, do as follows:

1. On MA5600T/MA5603T/MA5608T_S:

```
huawei(config)#undo ntp-service server disable
huawei(config)#ntp-service refclock-master 2
huawei(config)#vlan 2 standard
huawei(config)#port vlan 2 0/19 0
huawei(config)#interface vlanif 2
huawei(config-if-vlanif2)#ip address 10.10.10.10 24
huawei(config-if-vlanif2)#ntp-service manycast-server
huawei(config-if-vlanif2)#quit
```

2. On MA5600T/MA5603T/MA5608T_C:

```
huawei(config)#undo ntp-service server disable
huawei(config)#vlan 2 standard
huawei(config)#port vlan 2 0/19 0
huawei(config)#interface vlanif 2
huawei(config-if-vlanif2)#ip address 10.10.10.20 24
huawei(config-if-vlanif2)#ntp-service manycast-client
huawei(config-if-vlanif2)#quit
```

(IPv4) Assume the following configurations: MA5600T/MA5603T/MA5608T_S uses the local clock as the master NTP clock on stratum 2 and works in the NTP manycast mode, listening to clock synchronization packets that use 10:10::10:10/24 as the destination IP

address through IP address 10:10::10:10/24 of the L3 interface of VLAN 2 and establishing C/S connections with manycast clients, and MA5600T/MA5603T/MA5608T_C functions as the NTP manycast client, manycasting clock synchronization packets that use 10:10::10:10/24 as the destination IP address continousely through IP address 10:10::10:20/24 of the L3 interface of VLAN 2 and synchronizing with the clock on the manycast server by establishing C/S connections with the manycast server. To perform these configurations, do as follows:

1. On MA5600T/MA5603T/MA5608T_S:

```
huawei(config)#ipv6
huawei(config)#undo ntp-service ipv6 server disable
huawei(config)#ntp-service refclock-master 2
huawei(config)#vlan 2 standard
huawei(config)#port vlan 2 0/19 0
huawei(config)#interface vlanif 2
huawei(config-if-vlanif2)#ipv6 enable
huawei(config-if-vlanif2)#ipv6 address 10:10::10:10 24
huawei(config-if-vlanif2)#ntp-service manycast-server ipv6
huawei(config-if-vlanif2)#quit
```

2. On MA5600T/MA5603T/MA5608T_C:

```
huawei(config)#ipv6
huawei(config)#undo ntp-service ipv6 server disable
huawei(config)#vlan 2 standard
huawei(config)#port vlan 2 0/19 0
huawei(config)#interface vlanif 2
huawei(config-if-vlanif2)#ipv6 enable
huawei(config-if-vlanif2)#ipv6 address 10:10::10:20 24
huawei(config-if-vlanif2)#ntp-service manycast-client ipv6
huawei(config-if-vlanif2)#quit
```

(Optional) Configuring NTPv4 Control Signaling KOD

Kiss-o'-Death (KOD) packets are NTPv4-defined control signaling packets used for state advertisement during Network Time Protocol (NTP) message exchange and for access control. The KOD packets do not apply to time synchronization. The KOD packets can help the users to control the interval that the client transmits packets to the server.

Prerequisites

- The client sends NTP packets to the server for multiple times.
- The server filters packets sent by the client.

Context

KOD packets carry KOD codes that are indicated by a four-character ASCII character string. The KOD codes are called KISS codes. Different KISS codes have different meanings. The system supports RATE and DENY KISS codes.

KISS Code	Server Configuration	Application Scenario
DENY	Run the **ntp-service access** command to configure access control. Then the	This configuration applies when an NTP server is only configured for specified

KISS Code	Server Configuration	Application Scenario
	server sends KOD packets that carry DENY KISS codes to the client when the client's access request is denied because of restricted rights.	clients, which provides protection for the clients.
RATE	Run the **ntp-service access limited** command to configure rate control, and the **ntp-service discard** command to configure the minimum interval and average interval allowed by the server for the client to transmit packets. • When the client sends packets at an interval shorter than the minimum or average interval, the server sends KOD packets that carry RATE KISS codes to the client. • When the client sends packets at an interval longer than the minimum and average intervals, the server does not send any KOD packets.	This configuration applies to servers of poor performance.

Precaution

The following operations are performed only on servers.

Procedure

(Optional) In IPv6 systems, run the **ipv6** command to globally enable IPv6.

Step 1 Enable the NTP server functionality.

- Run the **undo ntp-service server disable** command to enable IPv4 NTP server functionality.
- Run the **undo ntp-service ipv6 server disable** command to enable IPv6 NTP server functionality.

Step 2 Run the **ntp-service kod-enable** command to enable the KOD function globally. Then the server sends KOD packets that carry appropriate KISS codes in different scenarios, which controls packet transmission of the client.

Step 3 Run the **ntp-service access** command to enable the access control function on the server. Then the server can control the client's access based on the client's rights. If the client's access is denied, the server sends KOD packets that carry **DENY** KISS codes.

Step 4 Run the **ntp-service discard** command to configure the minimum interval and average interval allowed by the server for the client to transmit packets. Then the server controls packet transmission of the client using **RATE** KISS codes in KOD packets based on packet transmission frequency.

----End

Example

Assume that server A in an IPv4 system has poor performance. The access control function needs to be enabled for server A to only receive NTP packets sent by clients at the minimum and average intervals longer than 3. To perform these configurations, do as follows:

```
huawei(config)#undo ntp-service server disable
huawei(config)#ntp-service kod-enable
huawei(config)#ntp-service access limited 3000
huawei(config)#ntp-service discard ntp-service discard min-interval 3 avg-interval 3
```

Assume that server B in an IPv6 system has poor performance. The access control function needs to be enabled for server B to only receive NTP packets sent by clients at the minimum and average intervals longer than 3. To perform these configurations, do as follows:

```
huawei(config)#ipv6
huawei(config)#undo ntp-service ipv6 server disable
huawei(config)#ntp-service kod-enable
huawei(config)#ntp-service access limited ipv6 3000
huawei(config)#ntp-service discard ntp-service discard min-interval 3 avg-interval 3
```

32.8.4 NTP Standards and Protocols Compliance

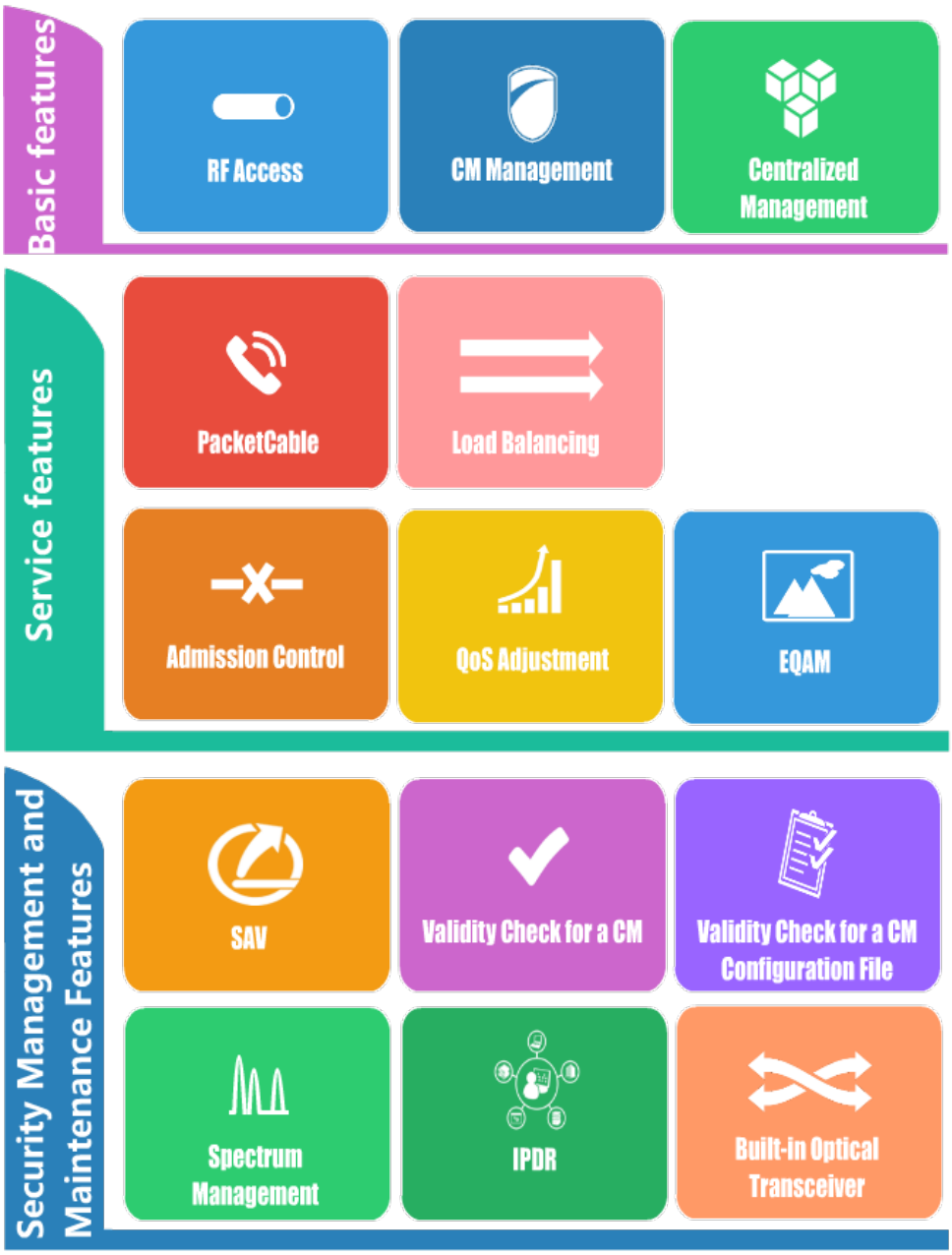
The following provides the reference documents of NTP:

- RFC 1305: Basis of the NTP module requirements specification
- RFC 5905: NTP version 4 Protocol and algorithm specification

33 D-CCAP

About This Chapter

Distributed converged cable access platform (D-CCAP) meets the requirements of triple-play network services over hybrid fiber coaxial (HFC) networks of multiservice operators (MSOs) because of unique advantages of the D-CCAPs, such as high bandwidth and supporting HFC networks.



33.1 D-CCAP Key Features and Usage Scenarios

Distributed converged cable access platforms (D-CCAPs) support the following features to meet various usage scenario requirements and help carriers flexibly provision services on cable networks.

Table 33-1 Basic D-CCAP features

Usage Scenario	Feature	Product Version
• Hybrid fiber coaxial (HFC) networks carry bidirectional	33.2 RF Access	This feature was introduced

Usage Scenario	Feature	Product Version
interactive broadband services based on CATV coaxial cable networks with the traditional analog transmission mode retained. • The HFC networks provides various services over existing coaxial cables. The services include telephony, broadcast television, VoD, and Internet access services.		in V800R013C00.
The D-CCAP manages cable modems (CMs) and implements service access and forwarding for the CMs. • CM registration process • CM status management • CM service flow forwarding	33.3 CM Management	This feature was introduced in V800R013C00.

Table 33-2 D-CCAP centralized management feature

Usage Scenario	Feature	Product Version
• Traditional cable modem termination systems (CMTSs) have the following disadvantages: – High access costs – Large upstream aggregation noises due to a great many of users connected to one radio frequency (RF) port – Lack of symmetric high bandwidth transmission due to a low upstream bandwidth • The MA5633 resolves traditional CMTS issues but introduces a new issue, higher operation expenditure (OPEX) for carriers' device management. More MA5633s are required for the same number of users	33.4 Centralized Management	This feature was introduced in V800R013C00.

Usage Scenario	Feature	Product Version
supported by a traditional CMTS because an MA5633 supports a smaller number of users. • To effectively manage MA5633s and reduce OPEX and total cost of operation (TCO), Huawei launches the D-CCAP centralized management feature.		

Table 33-3 D-CCAP service features

Usage Scenario	Feature	Product Version
VoIP services are required to be transmitted over existing cable networks.	33.5 PacketCable	This feature was introduced in V800R013C00.
Video data is required to be transmitted over cable networks based on the traditional IP network multicast technology.	DOCSIS Multicast	This feature was introduced in V800R013C10.
• Video data is encapsulated into UDP packets on IP networks, which cannot be transmitted on HFC networks. • Video data transmission on HFC networks uses existing cable television (CATV) coaxial cables on the HFC networks, thereby reducing network deployment costs.	33.7 EQAM-based Video Technologies	This feature was introduced in V800R015C00.
• System resources cannot meet the requirements for registering CMs or dynamically creating service flows. Services malfunction due to system resource exhaustion. • Bandwidths must be reserved for emergency calls to ensure the highest priority of the calls.	33.9 Admission Control	This feature was introduced in V800R013C00.

Table 33-4 D-CCAP networking features

Usage Scenario	Feature	Feature Description	Product Version	Command
Huawei has provided the solution of external optical transceiver+MA5633 for OOB applications. However, the external optical transceiver, an independent device, not only complicates networking but also increases installation and maintenance costs. To resolve the issues caused by external transceivers, Huawei has developed the MA5633 equipped with a built-in optical transceiver.	33.14 Built-in Optical Transceiver	A built-in optical transceiver consists of an optical receiver and an optical transmitter. • The optical receiver receives optical signals transmitted over optical fibers in the CATV transmission system, converts optical signals to electrical signals, and outputs RF signals. • The optical transmitter provides OOB signal backhaul for the VoD service. It uses the optical transmission backhaul module built in the MA5633 to send OOB data in upstream frequency bands to the head end video system for demodulation.	This feature was introduced in V800R015C00.	**optical agc range** *start-pos* **rf att** *value* **rf output switch** { **on** | **off** }

Table 33-5 D-CCAP security features

Vulnerability	Feature	Product Version
A malicious user forges the IP address of an authorized user and sends a great number of packets to attack the system. As a result, the system cannot process the services of authorized users.	33.11 SAV	This feature was introduced in V800R013C00.
Common Internet security threats include the following: • Unauthorized use: Resources are used without authorization. For example, attackers gain access to a computer system and use resources by guessing a user account and password combination. • Information theft: Attackers do not invade the target system, but sniff it to steal important data or information.	33.12 Validity Check for a CM	This feature was introduced in V800R013C00.
The following issues frequently occur on networks: • Issue 1: An unauthorized user tampers with the CM configuration file to obtain service resources that are different from the authorized ones. • Issue 2: An unauthorized user forges the CM configuration file of an authorized user to access a network and use network resources.	33.13 Validity Check for a CM Configuration File	This feature was introduced in V800R013C00.

Table 33-6 D-CCAP maintenance features

Usage Scenario	Feature	Feature Description	Product Version	Command
Huawei has provided the solution of	33.14 Built-in Optical	A built-in optical transceiver	This feature was introduced in	**optical agc range** *start-pos*

Usage Scenario	Feature	Feature Description	Product Version	Command
external optical transceiver+MA5633 for OOB applications. However, the external optical transceiver, an independent device, not only complicates networking but also increases installation and maintenance costs. To resolve the issues caused by external transceivers, Huawei has developed the MA5633 equipped with a built-in optical transceiver.	Transceiver	consists of an optical receiver and an optical transmitter. • The optical receiver receives optical signals transmitted over optical fibers in the CATV transmission system, converts optical signals to electrical signals, and outputs RF signals. • The optical transmitter provides OOB signal backhaul for the VoD service. It uses the optical transmission backhaul module built in the MA5633 to send OOB data in upstream frequency bands to the head end video system for demodulation.	V800R015C00.	**rf att** *value* **rf output switch** { **on** | **off** }
Noises must be minimized on upstream channels on HFC networks to	33.15 Spectrum Management	The spectrum management policy group specifies an adjustment rule	This feature was introduced in V800R013C10.	Frequency Spectrum Management

Usage Scenario	Feature	Feature Description	Product Version	Command
protect signal transmission of cable users' data and voice services.		and a set of parameters, including center frequencies, frequency bandwidths, and modulation profiles.		
The PON+D-CCAP solution has been widely used as a typical multiservice operator (MSO) solution, and its access users increment exponentially. The MSOs require an effective mechanism for collecting and reporting data for billing, fault location, and network running status monitoring to meet routine O&M requirements.	33.16 IPDR	The Internet Protocol Detail Record (IPDR) feature enables the MA5600T/MA5603T/MA5608T to collect data, encode the collected data in an IPDR-dedicated external data representation (XDR) format, and send the encoded data to an IPDR server. The collected data includes accounting information, running statuses of the cable modem (CM) served by the MA5600T/MA5603T/MA5608T, frequency spectrum information, debugging information, and CM termination system (CMTS) statistics. The IPDR feature	This feature was introduced in V800R013C00.	IPDR Configuration

Usage Scenario	Feature	Feature Description	Product Version	Command
		complies with data over cable service interface specification (DOCSIS) 2.0 and 3.0.		
The D-CCAP supports information query on remote CMs and queried data display. This facilitates CM management and fault location.	Remote information query	Remote information query supports periodic query and real-time query: • Periodic query: In centralized management mode, the D-CCAP automatically queries the data of all connected CMs within a period of time (configurable, with a default value of 30s). This prevents the D-CCAP from querying a large quantity of data within a short time. Each query duration is determined based on the number of connected CMs. • Real-time query: The	This feature was introduced in V800R013C10.	**display cable modem** { **all** | *cm-index* | *mac-address* | *ip-address* | *frameid/slotid/portid* [**upstream** *channel-id*] } **detail** **display cable modem** { *cm-index* | *mac-address* | *ip-address* } **remote-detail** **display cable modem phy** { *cm-index* | *mac-address* | *ip-address* | *frameid/slotid/portid*}

Usage Scenario	Feature	Feature Description	Product Version	Command
		D-CCAP queries the data of a single CM and obtains the CM data in real time.		
The MA5633 used in D-CCAP scenarios can rapidly identify an RF line where a noise source locates.	RF switch	An RF switch supports two statuses (**enable** and **disable**). The **enable** or **disable** status is used to work with the upstream spectrum scanning function provided by the U2000. In this way, the MA5633 can rapidly identify an RF line where a noise source locates, thereby improving troubleshooting efficiency.	This feature was introduced in V800R015C00.	**cable rf-out-switch** **display cable rf-out-switch**

33.2 RF Access

The radio frequency (RF) access feature enables the MA5600T/MA5603T/MA5608T to transmit cable television (CATV), voice, and Internet access services over existing coaxial cables through digital modulation.

33.2.1 Introduction

Hybrid fiber coaxial (HFC) networks have rich frequency band resources, wide coverage scope, and integrated service capabilities, which make them ideal candidates for carrying triple play services.

The RF access feature enables the HFC networks to carry interactive broadband services in both the downstream and upstream directions. HFC networks can therefore support conventional analog transmission modes and exploit existing CATV coaxial cable resources. This facilitates the integration of broadcast television, telecommunications, and Internet networks.

33.2.2 Principles

RF Parameters

Frequency

A center frequency and a frequency band determine a frequency range, within which packets are transmitted.

Modulation mode

Communication signals are classified as frequency modulation, phase modulation, and amplitude modulation by modulation mode. In digital signal transmission, commonly used modulation modes include quadrature amplitude modulation (QAM) and quadrature phase shift keying (QPSK).

Upstream channel type

An upstream channel can be of the A-TDMA, or S-CDMA type. A-TDMA is the acronym for advanced time division multiple access and S-TDMA is the acronym for synchronous code division multiple access.

MAC domain

A MAC domain is a logical sub-component of access device that is responsible for implementing all functions on a set of downstream channels and upstream channels. The access device serves all channels in a MAC domain. According to data over cable service interface specification (DOCSIS) requirements, a cable modem (CM) can access one or multiple upstream and downstream channels.

Mini-slot

A mini-slot is an upstream transmission unit. A group of mini-slots constitute an interval. Upstream channels use the time division multiplexing (TDM) technology to transmit data. An algorithm for allocating upstream channel bandwidths is used to set intervals for the upstream channels so that the upstream channels transmit data by timeslots.

DOCSIS Overview

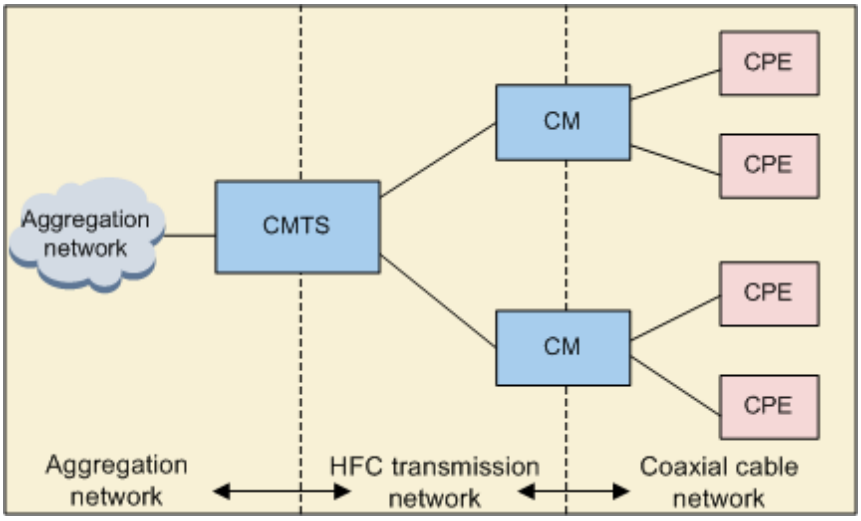
DOCSIS System Architecture

A DOCSIS system consists of an MA5600T/MA5603T/MA5608T, CMs, and an HFC transmission network.

The RF access feature enables:

- An RF port to connect to a coaxial cable network in the downstream direction.
- A PON or Ethernet network to connect to an aggregation network in the upstream direction.

Figure 33-1 DOCSIS system architecture



MA5600T/MA5603T/MA5608T

The MA5600T/MA5603T/MA5608T connects an aggregation network to an HFC network. It forwards network data, processes protocols, and modulates and demodulates RF signals. The MA5600T/MA5603T/MA5608T provides quality of service (QoS) required by CMs and allocates upstream bandwidth and service resources to the CMs based on CM requests and network QoS policies.

CM

A CM connects a customer premises equipment (CPE) device to an HFC network and uses DOCSIS protocols to transmit data in the HFC network.

HFC transmission network

An HFC transmission network uses both optical fibers and coaxial cables. HFC network signals consist of downstream signals and upstream signals.

- Downstream signals are CATV carriers or data carriers sent from the MA5600T/MA5603T/MA5608T to CMs in broadcast mode.
- Upstream signals are data carriers sent from CMs to the MA5600T/MA5603T/MA5608T in point to point (PTP) mode.

Channel Bonding

DOCSIS 3.0 supports channel bonding, which enables multiple channels to carry CM service flow data, thereby increasing the bandwidth of a given CM.

Channel bonding is supported in both the downstream and upstream directions:

- Downstream channel bonding: The MA5600T/MA5603T/MA5608T sends packets to a CM through multiple channels. The DOCSIS MAC header of each packet contains a packet sequence number. After receiving the packets, the CM restores the packets according to the packet sequence numbers.
- Upstream channel bonding: The CM fragments packets and sends them to the MA5600T/MA5603T/MA5608T using the timeslots authorized by the MA5600T/MA5603T/MA5608T. Each packet fragment carries a fragment sequence

number. After receiving the packet fragments, the MA5600T/MA5603T/MA5608T restores the packets according to the fragment sequence numbers.

The service flow data of a CM that supports channel bonding can be carried over a single downstream or upstream channel in a bonding group, or over multiple downstream or upstream channels. If the service flow data is carried over multiple downstream or upstream channels, the service flow is a bonded service flow and the channels carrying the service flow data constitute a channel bonding group.

The MA5600T/MA5603T/MA5608T supports dynamic bonding groups. After the channel bonding function is enabled, the MA5600T/MA5603T/MA5608T selects available channels from the channel bonding list supported by a CM to dynamically create a channel bonding group.

Upstream Channel Power Adjustment

During cable running, a CM's upstream TX power may not match the CM's upstream channel RX power due to some reasons, such as temperature changing or poor line quality. In this case, the CM will fail to go online. The upstream channel power adjustment feature enables the D-CCAP to control CMs' upstream TX power adjustment in a refined manner, thereby resolving the CM offline issue.

The process of adjusting a CM's upstream TX power and upstream channel RX power is as follows:

- When the CM goes online, it uses the minimum TX power or the upstream TX power stored when it went online the previous time as the current TX power and starts ranging.
- When receiving the CM's ranging request, the D-CCAP measures the upstream channel RX power of the CM. Then, the D-CCAP adjusts the CM TX power multiple times according to the pre-configured upstream channel RX power until the measured upstream channel RX power is within the permitted range.

Table 33-7 lists the operations required for adjusting the upstream channel RX power of a CM.

Table 33-7 Operations required for adjusting the upstream channel RX power of a CM

Criteria Parameter	Adjustment Activity	Purpose
|a - b| **NOTE** • a: specifies the upstream channel RX power of a CM. • b: specifies the CM's upstream channel RX power measured by the D-CCAP.	• If the value of |a - b| is less than or equal to the value of parameter **threshold**, specifying the threshold for adjusting the TX power, the D-CCAP does not send power adjustment messages to the CM. • If the value of |a - b| is greater than the value of parameter **threshold**, the D-CCAP sends a power adjustment message to the CM.	To prevent the issue of adjusting CM TX power because of the sudden or slight change of the attenuation between the CM and the D-CCAP due to environment changes (This issue adversely affects CM services.)
	• If the value of |a - b| is less than or equal to the value of parameter **continue**,	To prevent a CM online failure when the CM TX power reaches the maximum

Criteria Parameter	Adjustment Activity	Purpose
	specifying the threshold for continued ranging, the CM ranging succeeds. **NOTE** After the successful ranging, the D-CCAP also adjusts the CM TX power based on power adjustment parameter settings. • If the value of |a - b| is greater than the value of parameter **continue**, the CM continues ranging. If the number of consecutive ranging times exceeds the preset threshold, the ranging fails. **NOTE** During consecutive ranging, the D-CCAP also adjusts the CM TX power based on power adjustment parameter settings.	value
max (maximum power adjustment value each time)	The value of parameter **max** can be adjusted.	To prevent the required CM TX power after the adjustment from exceeding the maximum TX power supported by the CM

NOTE

- To configure upstream channel RX power adjustment parameters, run the **cable upstream** *channel-id* **power-adjust** { **threshold** *threshold* | **continue** *continue* | **max** *max* } * command.
- The relationships between the values of parameters **max**, **continue**, and **threshold** must comply with the formula (**max** ≥ **continue** ≥ **threshold**). The values of all the three parameters can be adjusted.
- If a CM fails to go online because of low TX power, increase the value of **continue** to enable the CM to go online.

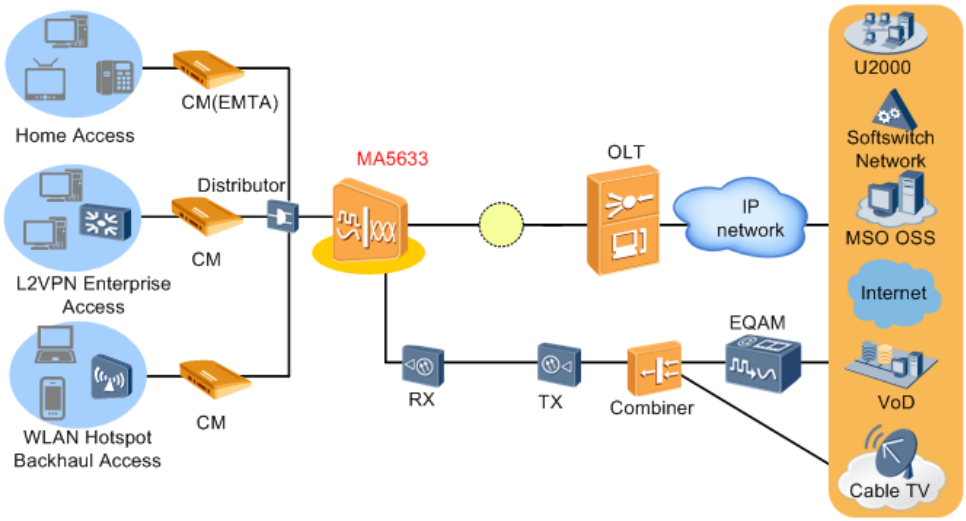
33.2.3 Application Scenarios

The D-CCAP networking scenarios support the following services:

- High-speed Internet (HSI), VoD, CATV, and PacketCable voice services for home users
- Wi-Fi, mobile radio backhaul, and L2VPN services for enterprises
- WLAN hotspot radio backhaul service by connecting access points (APs) to the MA5600T/MA5603T/MA5608T

Figure 33-2 shows the typical networking of the RF access feature.

Figure 33-2 Typical networking of the RF access feature



OLT: Optical Line Terminal	CM: Cable Modem
RX: Optical Receiver	TX: Optical Transmitter
EQAM: Edge Quadrature Amplitude Modulation	VoD: Video-On-Demand

33.2.4 Configuring RF Ports

This section describes how to configure RF ports on the MA5600T/MA5603T/MA5608T so that the MA5600T/MA5603T/MA5608T can provide the RF access feature.

Procedure

Run the **cable** command to configure RF port parameters.

- **freq-range**: indicates the upstream frequency range, which can be **european** or **north-american**.
- **mrc-mode** and **mtc-mode**: indicate the downstream and upstream channel bonding, respectively. After channel bonding is enabled, the MA5600T/MA5603T/MA5608T transmits the service flow data of a CM through two or more channels.
- **downstream annex**: indicates the downstream channel type, which can be **AnnexA** or **AnnexB**. Ensure that the downstream channel type and the upstream frequency range comply with the same standard.
  - If the downstream channel is of the AnnexA type, the upstream frequency must comply with the European standard. The frequency width of a single channel is 8 Mbit/s.
  - If the downstream channel is of the AnnexB type, the upstream frequency must comply with the North American standard. The frequency width of a single channel is 6 Mbit/s.

Step 1 **Optional:** Run the **display cable modulation-profile** command to query the information about an existing modulation profile for upstream channels.

A modulation profile defines parameters involved in signal processing for upstream channels.

Step 2 Run the **cable upstream** command to configure upstream RF parameters and activate an upstream channel.

- **frequency**: indicates the center frequency of a channel. Configure this parameter before activating a channel.
- **channel-width**: indicates the frequency width of a channel.
- **modulation-profile**: indicates the ID of the modulation profile used by the upstream channel. The profile defines the parameters used by the upstream channel during signal processing.
- **minislot-size**: a mini-slot is an upstream transmission unit. A group of mini-slots constitute an interval, and a group of ticks constitute a mini-slot. Upstream channels use the time division multiplexing (TDM) technology to transmit data. An algorithm for allocating upstream channel bandwidths is used to set intervals for the upstream channels so that the upstream channels transmit data by timeslots.
- **rf-power**: indicates the upstream channel receive power.

The center frequency and frequency width determine a frequency range for packet transmission. For example, if a channel complies with the European standard, the center frequency is 55 MHz, and the frequency width is 6400 kHz, then the channel frequency ranges from 51.8 MHz to 58.2 MHz.

Step 3 Run the **cable downstream** command to configure downstream RF parameters and activate a downstream channel.

- **frequency**: indicates the center frequency of a channel. Configure this parameter before activating a channel.
- **modulation**: indicates the modulation mode of a downstream channel. A greater modulation mode increases the channel bandwidth but reduces the anti-interference capability. Set this parameter based on site requirements. The default modulation mode is QAM 256.
- **interleave-depth**: indicate the downstream channel interleaving depth. This parameter is available only when the channel type complies with the North American standard.
- **rf-power**: indicates the downstream channel transmit power. Set this parameter to a proper value based on site requirements.

----End

Example

The following is an example of the configurations used to configure RF port 1/1/0 on the MA5600T/MA5603T/MA5608T:

- The interval in which the MA5600T/MA5603T/MA5608T sends the MAC domain description to CMs through the downstream channel is 1800 ms.
- The upstream frequency complies with the European standard.
- Channel bonding is enabled in both the downstream and upstream directions.
- The downstream channel is of the AnnexA type.
- The center frequency of the downstream channel is 440.00 MHz.
- The center frequency of the upstream channel is 10.00 MHz.
- Other parameters use default settings.

```
huawei(config-if-cable-1/1/0)#cable mdd-interval 1800 freq-range european mrc-mode
enable mtc-mode e
nable downstream annex annexA
huawei(config-if-cable-1/1/0)#cable upstream 1 frequency 10.00 enable
huawei(config-if-cable-1/1/0)#cable downstream 1 frequency 440.00 enable
```

33.2.5 Standards and Protocols Compliance

The RF access feature complies with the following standards and protocols:

- CM-TR-OSSIv3.0-CM-V01-08092
- CM-TR-MGMTv3.0-DIFF-V01-071228
- CM-SP-SECv3.0-I13-100611
- CM-SP-PHYv3.0-I09-101008
- CM-SP-OSSIv3.0-I14-110210
- CM-SP-MULPIv3.0-I15-110210
- CM-SP-DRFI-I11-110210

33.3 CM Management

A cable modem (CM) uses the data over cable service interface specification (DOCSIS) protocol to connect customer premises equipment (CPE) devices to a carrier's hybrid fiber coaxial (HFC) network. The MA5600T/MA5603T/MA5608T manages CMs as well as CM registration, CM service flow forwarding, CM configuration file parsing, and CM status management.

What Is CM Management

Table 33-8 lists CM processes managed by the MA5600T/MA5603T/MA5608T.

Table 33-8 CM processes managed by the MA5600T/MA5603T/MA5608T

Item	Description
CM registration	CM registration is a process spanning from the time when a CM is powered on to the time when the CM is ready for provisioning services for users.
CM service flow forwarding	After receiving packets from a CPE, a CM sends the packets to the optical line terminal (OLT) through a radio frequency (RF) port. Then, the OLT processes the packets and sends them to the upper-layer network through an uplink port, implementing service flow forwarding on the access side.
CM status management	The status of a CM can be queried on the MA5600T/MA5603T/MA5608T for fault location purposes.
CM configuration file parsing	Users manage HFC networks and configure services through a configuration file. Using the information in the configuration file, the OLT and CMs can connect to a Trivial File Transfer Protocol (TFTP) server and the MA5600T/MA5603T/MA5608T can manage CMs and

Item	Description
	issue services to the CMs.

33.3.2 Principles of CM Management

CM Registration

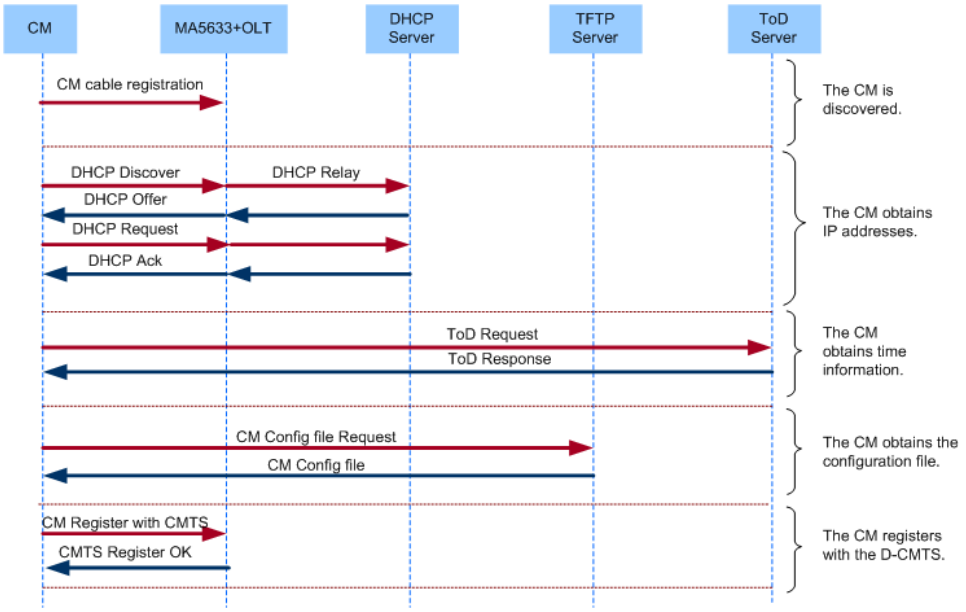
CM registration is a process spanning from the time when a CM is powered on to the time when the CM is ready for provisioning services for users.

Before powering on a CM, ensure that the following requirements are met:

- The CM has been configured on a DHCP server. The configuration information includes the IP address of the CM, gateway address, IP address of the time of day (ToD) server, configuration file name, and IP address of the TFTP server where the configuration file is stored.
- The configuration file is available and has been saved in a specified path on the TFTP server.
- The OLT has been configured, including security authentication and service VLANs (S-VLANs) configuration.

Figure 33-3 shows the CM registration process.

Figure 33-3 CM registration process



DHCP server: provides the IP, gateway, TFTP server, and ToD server addresses for a CM.

TFTP server: stores CM configuration files. A CM obtains a configuration file from the TFTP server.

ToD server: provides the date and time for a CM. After obtaining this information, the CM reports event logs with accurate date and time information. This facilitates device management.

The following table describes the CM registration process.

Stage	Procedure	Remarks
The CM is discovered.	The CM sets up the temporary connection with the MA5600T/MA5603T/MA5608T. 1. The CM is powered on. 2. The CM selects upstream channels, scans downstream channels, and locks the main downstream channel. The CM transmits all packets over the upstream and downstream channels. 3. The CM starts ranging and enters the automatic discovery stage. The MA5600T/MA5603T/MA5608T creates a temporary service flow for the CM. 4. (Optional) Enable EAE authentication.	None
The CM obtains its IP address and the IP addresses of the TFTP and ToD servers.	1. The CM initiates a DHCP request. 2. The MA5600T/MA5603T/MA5608T service module captures the DHCP packet, adds the MAC address and physical port number of the CM to the DHCP Option 82 field in the packet, performs DHCP relay, and forwards the DHCP packet to the DHCP server through an uplink port. 3. The DHCP server checks CM configurations according to the MAC address of the CM, allocates an IP address to the CM, and uses the DHCP Option 82 field to send the CM configuration file name and the IP addresses of the TFTP and ToD servers to the MA5600T/MA5603T/MA5608T. 4. The MA5600T/MA5603T/MA5608T service module captures the DHCP response packet and learns the mapping between the MAC address of the CM and the configuration file name. This information is subsequently used when the CM requests the configuration file from the TFTP server.	None
The CM obtains	1. The CM initiates a ToD request.	Based on the date and

Stage	Procedure	Remarks
time information.	2. The ToD server responds and the CM obtains the date and time information.	time information obtained by the CM, users can manage devices and obtain device running information.
The CM obtains the configuration file.	1. The CM requests the configuration file from the TFTP server according to the configuration file name contained in the DHCP Option 82 field. 2. The TFTP server sends the configuration file to the CM in TFTP mode.	A CM configuration file defines service flows, quality of service (QoS), and security policies. After the CM obtains a configuration file, it initiates a registration request to the MA5600T/MA5603T/MA5608T.
The CM registers with the MA5600T/MA5603T/MA5608T.	1. After the CM parses the configuration file, it initiates a registration request to the MA5600T/MA5603T/MA5608T. The registration request is in the type-length-value (TLV) format and contains service flow parameters. 2. The MA5600T/MA5603T/MA5608T performs a message integrity check (MIC) on the CM registration request to prevent the CM from modifying the configuration file without authorization. For details about the MIC, see **MIC** in 33.13 Validity Check for a CM Configuration File. 3. The MA5600T/MA5603T/MA5608T performs an X.509 certificate authentication on the CM. For details about the authentication, see **X.509 Authentication** in 33.13 Validity Check for a CM Configuration File. 4. The MA5600T/MA5603T/MA5608T checks service parameter settings and resources. It allocates service flow resources to the CM according to the parameter settings in the CM configuration file only if the service parameter settings comply with the configuration file and the remaining service resources meet the resource request requirement. In addition, the MA5600T/MA5603T/MA5608T maps the service flow resources to DOCSIS service flows. 5. The CM successfully registers with the MA5600T/MA5603T/MA5608T.	None

Stage	Procedure	Remarks
	CM service flows can be forwarded.	

CM Service Flow Forwarding

After receiving packets from a CPE, a CM sends the packets to the OLT through an RF port. Then, the OLT processes the packets and sends them to the upper-layer network through an uplink port, implementing service flow forwarding on the access side.

Figure 33-4 shows the processes of forwarding CM service flows in the upstream and downstream directions.

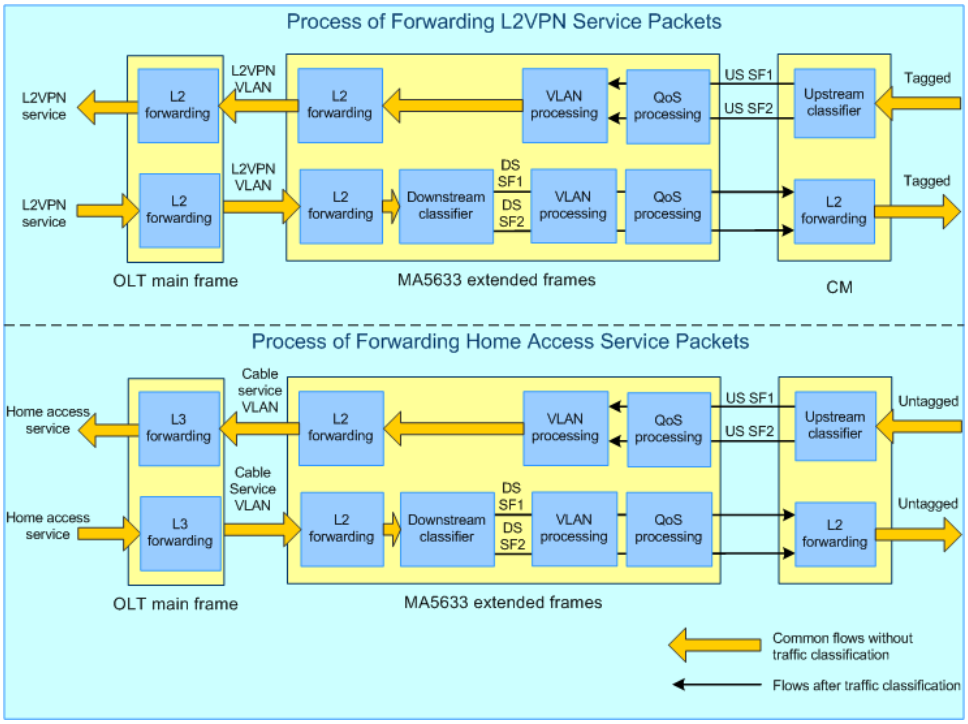
In the upstream direction

1. After receiving packets from CPEs, the CM classifies service flows and sends the packets to the MA5600T/MA5603T/MA5608T over the classified service flows. The packets from the home service are untagged and the packets from the Layer 2 virtual private network (L2VPN) service are tagged.
2. The MA5600T/MA5603T/MA5608T performs QoS operations, including bandwidth management, scheduling policy configuration, and priority remarking, based on service flows.
3. The MA5600T/MA5603T/MA5608T adds a VLAN tag to the packets based on service flows.
4. The MA5600T/MA5603T/MA5608T Layer 2 forwarding module forwards the packets at Layer 2 to an uplink port and then to the OLT using the VLAN ID and CM MAC address.
5. The OLT Layer 3 forwarding module forwards home service packets at Layer 3 to the upper-layer network.
6. The OLT Layer 2 forwarding module forwards L2VPN service packets at Layer 2 to the upper-layer network.

In the downstream direction

1. The OLT Layer 3 forwarding module forwards home service packets at Layer 3 to the MA5633 extended frame.
2. The OLT Layer 2 forwarding module forwards L2VPN service packets at Layer 2 to the MA5633 extended frame.
3. The MA5600T/MA5603T/MA5608T Layer 2 forwarding module identifies the CM using the VLAN ID and CM MAC address.
4. The MA5600T/MA5603T/MA5608T classifies traffic on common downstream service flows so that the packets can be transmitted over multiple downstream service flows by service type.
5. The MA5600T/MA5603T/MA5608T removes the VLAN tag from the packets based on service flows.
6. The MA5600T/MA5603T/MA5608T performs QoS operations, including bandwidth management, scheduling policy configuration, and priority remarking, based on service flows.
7. The CM forwards the packets to CPEs using the CPE MAC address.

Figure 33-4 CM service flow forwarding



CM Configuration File Parsing

Users manage HFC networks and configure services through a configuration file. Using the information in the configuration file, the MA5600T/MA5603T/MA5608T and CMs can connect to a TFTP server and the MA5600T/MA5603T/MA5608T can manage CMs and issue services to the CMs.

- Contents of the configuration file: The configuration file contains the global control, traffic classification, and QoS parameters of a CM.
- Format of the configuration file: A CM configuration file is written in binary in the TLV format.
  - Type: an 8-byte identifier that defines the parameter name.
  - Length: an 8-byte identifier that defines the value field length.
  - Value: an identifier containing 1-254 bytes that defines a parameter value.
- Editor for the configuration file: The CableLabs Config File Editor, released by CableLabs, is used to edit configuration files.
- Issuing mode of the configuration file: A CM obtains a configuration file from the TFTP server. When the CM registers with the MA5600T/MA5603T/MA5608T, the MA5600T/MA5603T/MA5608T creates service flows according to parameter settings in the configuration file.

The process of parsing a configuration file is as follows:

1. The MA5600T/MA5603T/MA5608T parses the CM configuration file.

   The MA5600T/MA5603T/MA5608T checks the validity and completeness of configuration file parameters, (for example, flow classification parameters) according to *CM-SP-MULPIv3_0-I09-090121*. If the check results do not meet the

CM-SP-MULPIv3_0-I09-090121 requirements, the MA5600T/MA5603T/MA5608T forces the CM to go offline.

2. The MA5600T/MA5603T/MA5608T loads the configurations defined in the configuration file.

   After the MA5600T/MA5603T/MA5608T parses the CM configuration file, it creates DOCSIS service flows between itself and the CM according to the service configurations in the configuration file.

   If the MA5600T/MA5603T/MA5608T successfully creates the DOCSIS service flows, the CM goes online. Otherwise, the MA5600T/MA5603T/MA5608T forces the CM to go offline.

33.3.3 Configuring CM Management

This section describes how to configure CM management.

Configuring a CM S-VLAN

A CM S-VLAN identifies CM service flows. All service flows (excluding L2VPN service flows, but including the CM management packet flows that carry the L2VPN service) of a CM use an S-VLAN ID. After a CM S-VLAN is configured, the MA5600T/MA5603T/MA5608T automatically creates service flows for the CM when the CM goes online.

Prerequisites

The CM S-VLAN to be configured has been created by running the **vlan** command.

Context

An L2VPN S-VLAN can be configured through a CM configuration file or commands. Only one mode is used each time generally. If both modes are used, the S-VLAN configuration through commands preferentially takes effect on the MA5600T/MA5603T/MA5608T when the CM goes online.

- S-VLAN configuration through a CM configuration file: Run the **port vlan** command to add the uplink port for the CM to the S-VLAN. The L2VPN S-VLAN is specified in the CM configuration file. When the CM goes online, the MA5600T/MA5603T/MA5608T obtains the L2VPN S-VLAN from the CM configuration file and automatically creates L2VPN service flows.
- S-VLAN configuration through commands: Manually configure the L2VPN S-VLAN on the MA5600T/MA5603T/MA5608T. For details, see the following procedure.

Procedure

- Configure a home access S-VLAN or CM management VLAN.
  a. Run the **cable service-vlan** command to configure a CM S-VLAN.

     The default S-VLAN ID is 1. The S-VLAN can only be modified.

     A newly configured S-VLAN takes effect after the CM goes online the next time. To make the S-VLAN immediately take effect, run the **cable modem reset all** command to restart all CMs.

NOTICE

Exercise caution when running this command because the execution interrupts services.

b. Run the **port vlan** command to add the uplink port to the S-VLAN.

- Configure an L2VPN S-VLAN through commands.
  a. Run the **cable l2vpn dot1q** command to configure an L2VPN S-VLAN.

     The L2VPN S-VLAN must be an existing VLAN of the smart type and with the common attribute. In addition, the L2VPN S-VLAN cannot be a reserved VLAN.

  b. Run the **port vlan** command to add the uplink port to the S-VLAN.

----End

Example

To configure CM S-VLAN 2 so that all of its service flows use VLAN ID 2, run the following commands:

```
huawei(config)#cable service-vlan 2
huawei(config)#port vlan 2 0/19 0
```

To configure L2VPN S-VLAN 2 for the CM with MAC address 0000-0000-1111 that connects to port 1/1/0, run the following commands:

```
huawei(config)#cable l2vpn dot1q 1/1/0 cm 0000-0000-1111 2
huawei(config)#port vlan 2 0/19 0
```

Configuring the Automatic CM Service VLAN Allocation Function

This section describes how to configure the automatic cable modem (CM) service VLAN allocation function to simplify data configurations if the following requirements are met: Layer 3 services need to be provisioned for home users; the gateway is planned on an MA5600T/MA5603T/MA5608T; service VLANs are not planned for CMs. Then, the MA5600T/MA5603T/MA5608T automatically allocates an idle service VLAN to the CMs.

Prerequisites

The MA5600T/MA5603T/MA5608T functions as a Layer 3 gateway.

Procedure

Run the **cable bundle reserved vlan** command to configure the CM service VLAN range reserved on the MA5600T/MA5603T/MA5608T.

A CM service VLAN cannot be a VLAN reserved in the system or an existing VLAN. To query reserved VLANs, run the **display vlan reserve** command. To query an existing VLAN, run the **display vlan** command.

Step 1 Run the **vlan** command to create a global CM service VLAN.

The global CM service VLAN must be a super VLAN.

Step 2 Run the **interface vlanif** command to create a Layer 3 interface.

Step 3 Run the **vlan service-profile** command to create a VLAN service profile.

Step 4 Run the **cable bundle** command to create a cable bundle.

Step 5 Run the **cable bundle interface** command to configure the Layer 3 interface of the cable bundle.

Step 6 Run the **cable bundle service-profile** command to bind the VLAN service profile to the service VLANs that can be automatically allocated to the CMs in the cable bundle.

Step 7 Run the **cable bundle member** command to add a CM to the cable bundle.

----End

Example

The following configurations are used as an example to configure the automatic CM service VLAN allocation function on the basis that the MA5600T/MA5603T/MA5608T works as a Layer 3 gateway.

Table 33-9 Data plan

Item	Data
IDs of the VLANs reserved for the cable bundle	500-600
Cable bundle group	Name: area • ID of the service VLAN: 100; type: super VLAN • VLAN service profile: profile1 • Cable bundle members: 1/1, 2/1, 3/1, and 4/1

```
huawei(config)#cable bundle reserved vlan add 500-600
huawei(config)#vlan 100 super
huawei(config)#interface vlanif 100
huawei(config-if-vlanif100)#quit
huawei(config)#vlan service-profile profile-name profile1
huawei(config-vlan-srvprof-1)#quit
huawei(config)#cable bundle area
huawei(config-cable-bundle-area)#cable bundle interface vlanif 100
huawei(config-cable-bundle-area)#cable bundle service-profile profile-name profile1
huawei(config-cable-bundle-area)#cable bundle member 1/1
huawei(config-cable-bundle-area)#cable bundle member 2/1
huawei(config-cable-bundle-area)#cable bundle member 3/1
huawei(config-cable-bundle-area)#cable bundle member 4/1
```

Configuring DOCSIS Event Reporting

DOCSIS 3.0 defines various DOCSIS events, such as authentication failure and CM certificate error events. Configuring DOCSIS event reporting helps you obtain MA5600T/MA5603T/MA5608T and CM running status.

Procedure

Run the **cable event loghost ip***ServerIpAddress* command to set the IP address of a log host to which DOCSIS events are reported.

Step 1 Run the **cable event level report** { **localVolatile** | **trap** | **syslog** } command to configure the mode and priority of reporting DOCSIS events.

The modes of reporting DOCSIS events include **localVolatile**, **trap**, and **syslog**.

- **localVolatile**: indicates that the DOCSIS events are saved to MA5600T/MA5603T/MA5608T local logs but not to the buffer. The MA5600T/MA5603T/MA5608T does not save the DOCSIS events and clears them after being reset. Configure the **localVolatile** mode before setting the mode of reporting DOCSIS events to **trap** or **syslog**.
- **trap**: reports DOCSIS events to a log host, which is generally the U2000, as traps. After setting this mode, run the **snmp-agent target-host trap-hostname** command to configure the log host.
- **syslog**: reports DOCSIS events to a log host as system logs, which are displayed after the log host software parses them.

The mode of reporting DOCSIS events varies depending on the priorities of the DOCSIS events.

- Default mode of reporting DOCSIS events with emergency or alert priorities: **localVolatile**
- Default mode of reporting DOCSIS events with critical, error, warning, or notice priorities: **trap**, **syslog**, and **localVolatile**
- Default mode of reporting DOCSIS events with information or debugging priorities: not reported

Step 2 (Optional)Run the **cable cm-status event** command to set parameters for a CM to report events.

----End

Example

The following is an example of the configurations used to enable the CM management feature:

- Priority of DOCSIS events: alert
- Mode of reporting DOCSIS events to the U2000: **trap**
- Mode of reporting DOCSIS events to the log host: **syslog**
- IP address of the U2000: 10.10.10.10
- IP address of the log host: 10.10.20.10

```
huawei(config)#cable event alert report trap
huawei(config)#snmp-agent target-host trap-hostname huawei address 10.10.10.10
trap-paramsname docsis
huawei(config)#cable event alert report syslog
huawei(config)#cable event loghost ip 10.10.20.10
```

33.3.4 CM Management Reference Files

The CM management feature complies with the following standards:

- CM-TR-OSSIv3.0-CM-V01-08092
- CM-TR-MGMTv3.0-DIFF-V01-071228
- CM-SP-SECv3.0-I13-100611
- CM-SP-PHYv3.0-I09-101008
- CM-SP-OSSIv3.0-I14-110210
- CM-SP-MULPIv3.0-I15-110210
- CM-SP-DRFI-I11-110210

33.4 Centralized Management

Traditional cable modem termination systems (CMTSs) have some disadvantages and the distributed converged cable access platform (D-CCAP) solution resolves CMTS issues and therefore is attracting more and more attention from carriers. However, an MA5633 supports a smaller number of users than a traditional CMTS and more MA5633s are required for the same number of users. If the MA5633s are managed as independent NEs, the operation expenditure (OPEX) for carriers' device management is high. To effectively manage MA5633s and reduce OPEX and total cost of ownership (TCO), Huawei develops the D-CCAP centralized management feature.

Introduction

D-CCAP centralized management: An optical line terminal (OLT) deployed in a branch equipment room functions as a main frame and connects to MA5633s (MA5633s, extended frames) located on optical nodes by using cascading boards. In this manner, remote extended frames are no longer standalone NEs (no longer allocated independent management IP addresses), but are managed by the main frame. These remote extended frames are regarded as remote service boards of the main frame and have the same functions and features as those of the main frame. The OLT and the MA5633 can be used as a traditional CMTS. The advantages of the MA5633 centralized management feature are as follows:

- This feature is compatible with the existing network.

  A multiservice operator (MSO) network consists of a coaxial cable network on the user side, a hybrid fiber coaxial (HFC) transmission network, and a metro aggregation network. In MA5633 centralized management, the OLT and the MA5633s are compatible with the existing MSO network and can replace traditional CMTSs. The coaxial cable network on the user side and the metro aggregation network remain unchanged. The MA5633 connects to a cable modem (CM) through a radio frequency (RF) port.
- This feature supports more users without increasing the number of management objects.
- Software commissioning is not required for remote extended frames and one onsite operation suffices.
- Service provisioning interfaces on remote extended frames remain unchanged. The user interface for extended frames is the same as that for a service board of the main frame, enabling consistent user experience.

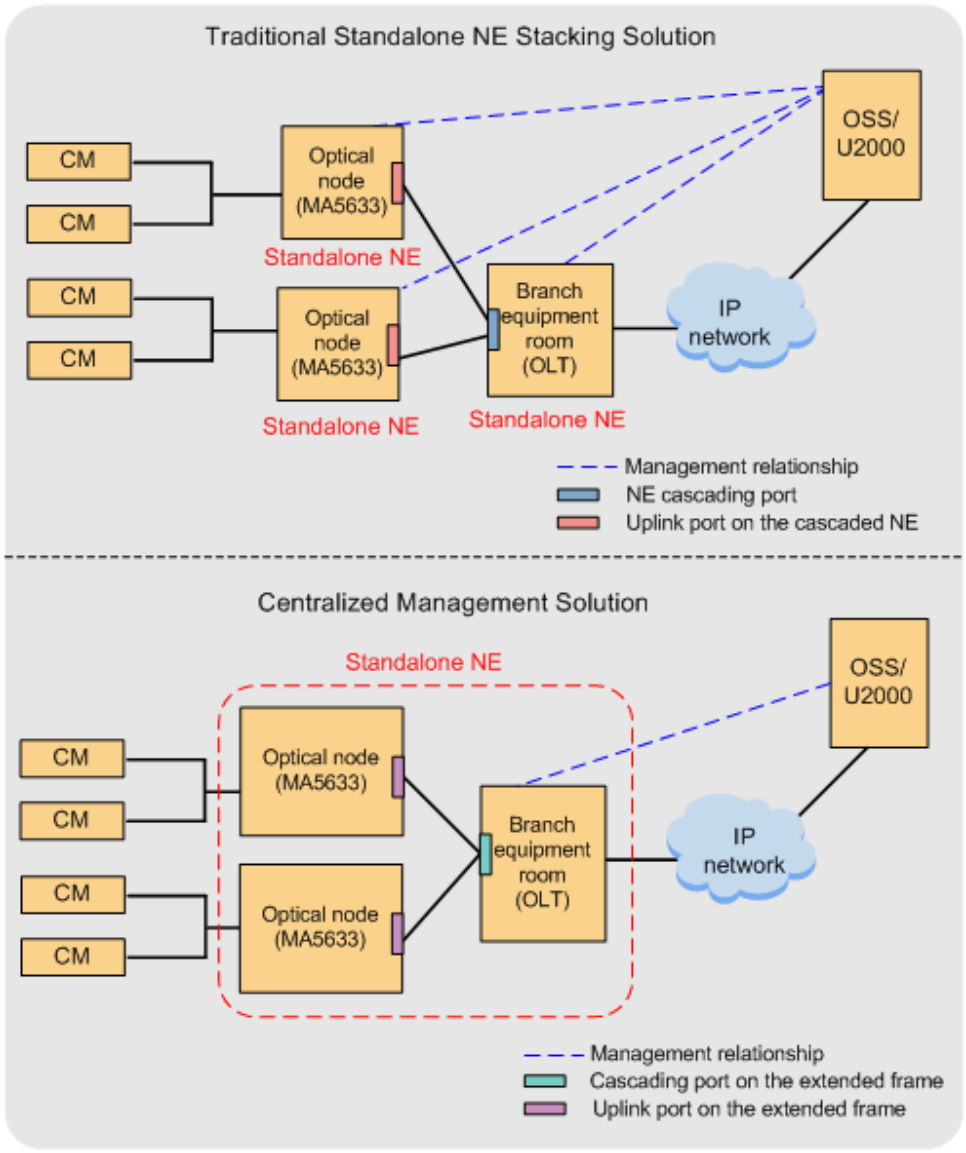
  Each MA5633 deployed at the remote end can be regarded as a new service board deployed on the main frame. The MA5633 does not need to interconnect with the

upper-layer OSS or NMS system, reducing operating expense (OPEX) and total cost of ownership (TCO) for carriers.

- PON ports and Layer 2 forwarding operations are transparent between the OLT and the MA5633s, simplifying configurations for carriers.
- A new MA5633 after replacement does not require manual configurations because the OLT automatically issues configurations for it.

Figure 33-5 shows the differences between the centralized management solution and the traditional standalone NE stacking solution.

Figure 33-5 Comparison between the two solutions



33.4.2 Basic Concepts

System Management Policy

An optical line terminal (OLT) supports the configuration of the system management policy (**extend-frame** or **stand-alone**) by running the **sysman centralized-mgmt primary** command. The system management policy defines the management mode for the MA5633s to function as extended frames or standalone NEs.

- If the system management policy is set to **extend-frame**, the MA5633s initiate a registration request to the OLT as extended frames.
- If the system management policy is set to **stand-alone**, the MA5633s initiate a registration request to the OLT as standalone NEs.

System Management Mode Autonegotiation

An MA5633 can function as an extended frame of an OLT or a standalone NE. Accordingly, the MA5633 software supports the extended frame and standalone NE modes, which are configured before delivery. The software takes effect in the extended frame mode by default.

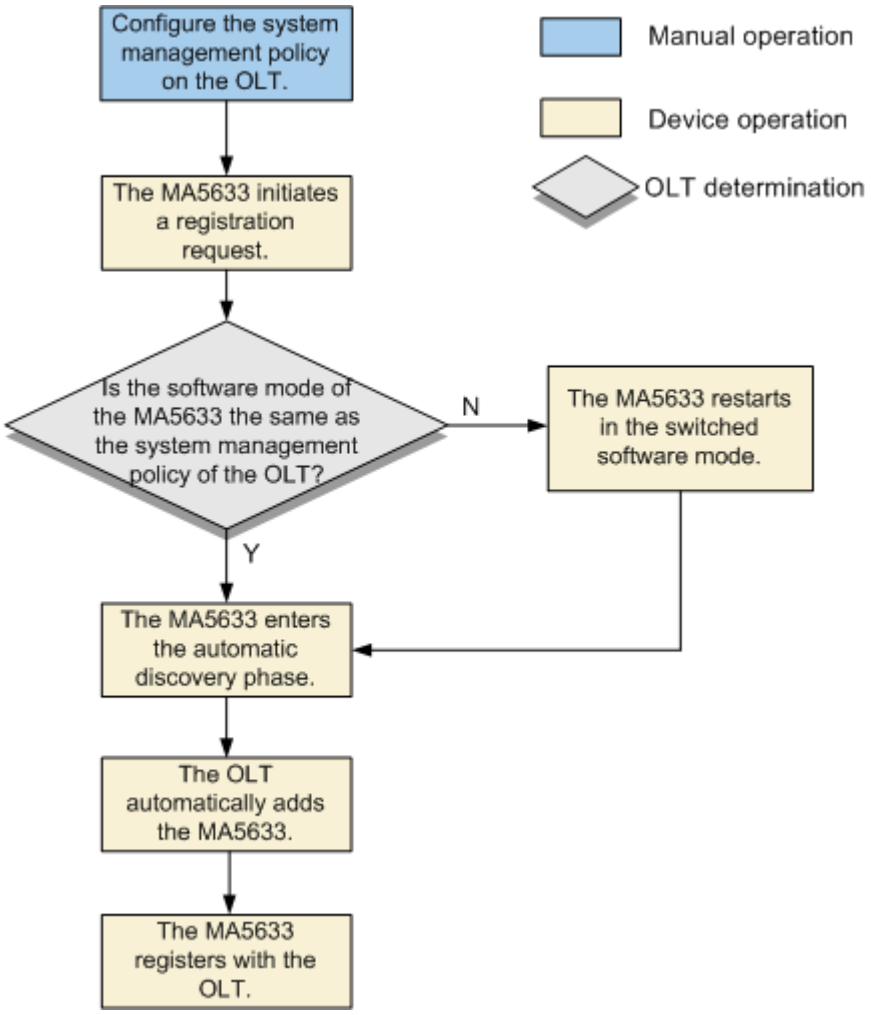
Upon power-on, the MA5633 negotiates with the OLT about the system management policy. This ensures that the software mode of the MA5633 is the same as the system management mode of the OLT.

- In GPON upstream transmission, the MA5633 uses the extended optical network terminal management and control interface (OMCI) proprietary protocol for negotiation.
- In GE upstream transmission, the MA5633 uses proprietary bridge protocol data unit (BPDU) packets for negotiation.

If the software mode of the MA5633 is different from the system management mode of the OLT, the OLT notifies the MA5633 of switching the software mode. Then, the MA5633 restarts in the new software mode.

Figure 33-6 shows the flowchart for autonegotiation of the system management mode.

Figure 33-6 Flowchart for autonegotiation of the system management mode



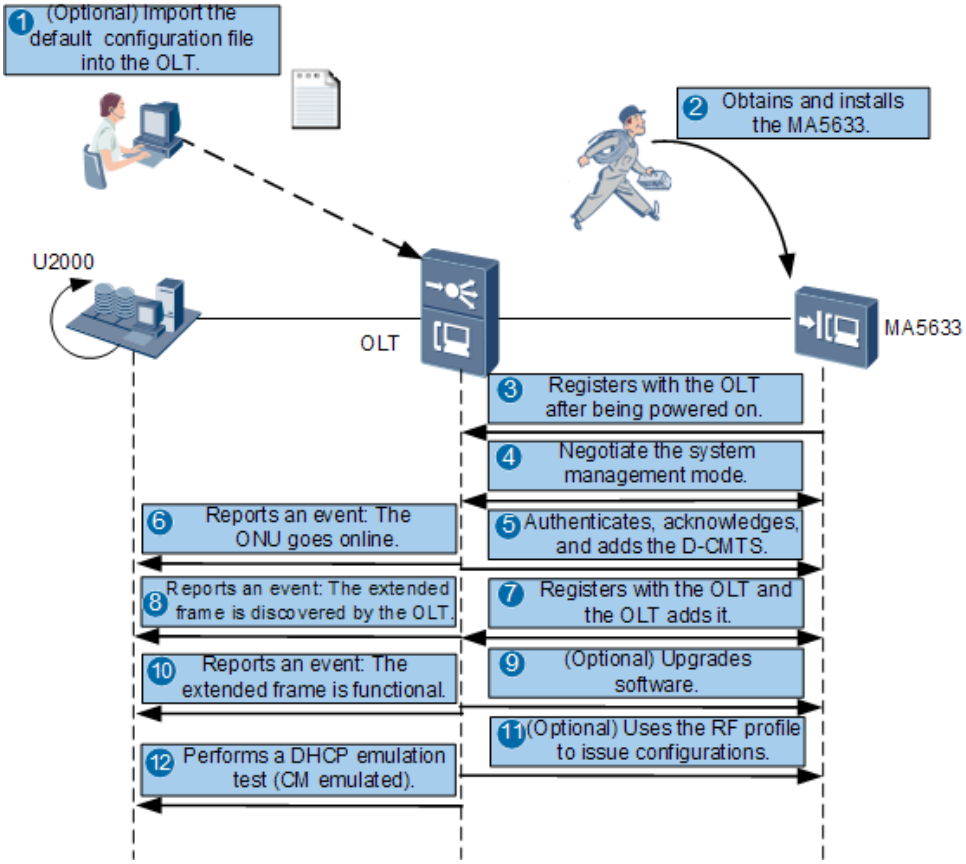
33.4.3 Centralized Management for Remote GPON Extended Frames

Automatic Discovery for Remote GPON Extended Frames

The automatic discovery for remote GPON extended frames enables an optical line terminal (OLT) to automatically acknowledge and add a remote GPON extended frame. This simplifies the flows for ONU authentication, and extended frame registration and addition. This function applies to the scenario where a great number of distributed cable modem termination systems (MA5633s) need to be deployed.

Figure 33-7 shows the process of deploying automatic discovery for remote GPON extended frames.

Figure 33-7 Process of deploying automatic discovery for remote GPON extended frames



1. **Optional:** The software commissioning personnel import the default configuration file, which includes RF parameter settings of the MA5633, into the OLT.
2. The hardware installation personnel obtain the data planning table, fetch the MA5633 device according to the MAC address in the data planning table, and install the MA5633 device on site.
3. The hardware installation personnel power on the MA5633 device and make optical paths available. The MA5633 device initiates a registration request to the OLT.
4. The MA5633 device and the OLT negotiate the system management mode. For details, see **System Management Mode Autonegotiation** in 33.4.2 Basic Concepts.
5. The OLT automatically authenticates, acknowledges, and adds the MA5633 as an optical network unit (ONU).
6. The OLT reports an event to the U2000, indicating that the MA5633 device has been online.
7. The MA5633 device initiates a registration request to the OLT and the OLT adds the MA5633 device as an extended frame.
8. The OLT reports an event to the U2000, indicating that the OLT has automatically discovered the MA5633 device.
9. **Optional:** The OLT automatically upgrades the MA5633 software if the MA5633 software version is different that the OLT software version.
10. The OLT reports an event to the U2000, indicating that the MA5633 device is functioning properly.

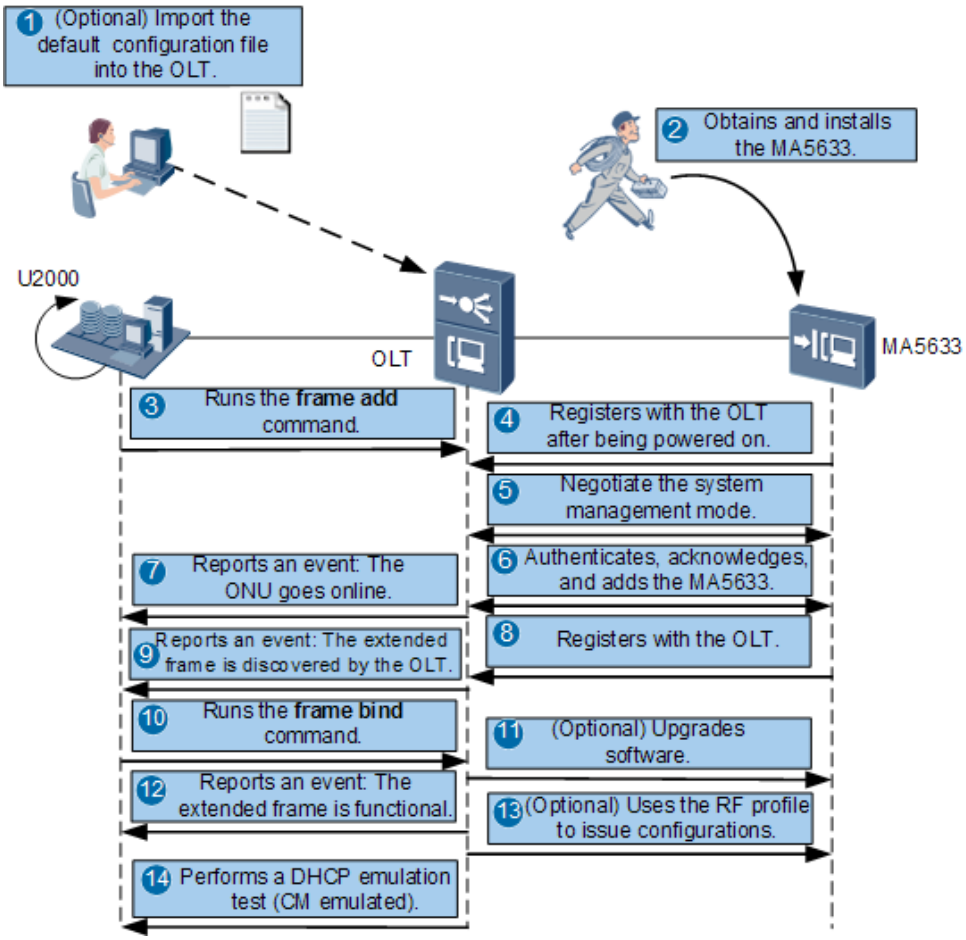
11. **Optional:** If a radio frequency (RF) line initialization profile has been configured on the OLT, the OLT automatically issues RF line configurations to the MA5633 device according to the profile after the OLT adds the MA5633 device.
12. The software commissioning personnel initiate a Dynamic Host Configuration Protocol (DHCP) emulation test through the U2000 to verify the data configuration. Then, the data configuration is complete.

Remote GPON Extended Frame Deployment in Offline Mode

The remote GPON extended frame deployment in offline mode enables the OLT to automatically acknowledge and add a remote GPON extended frame. This simplifies the flows for ONU authentication. This function facilitates device management and applies to the scenario where MA5633s need to be deployed one by one.

Figure 33-8 shows the process of deploying a remote GPON extended frame in offline mode.

Figure 33-8 Process of deploying a remote GPON extended frame in offline mode



1. **Optional:** The software commissioning personnel import the default configuration file, which includes RF parameter settings of the MA5633, into the OLT.
2. The hardware installation personnel obtain the data planning table, fetch the MA5633 device according to the MAC address in the data planning table, and install the MA5633 device on site.

3. The software commissioning personnel run the **frame add** command on the OLT to add the MA5633 device in offline mode.
4. The hardware installation personnel power on the MA5633 device, make optical paths available, and send the serial number (SN) of the MA5633 device to the U2000.
5. The MA5633 device and the OLT negotiate the system management mode to ensure that the software mode of the MA5633 is the same as the system management mode of the OLT. For details, see **System Management Mode Autonegotiation** in 33.4.2 Basic Concepts.
6. The OLT automatically authenticates, acknowledges, and adds the MA5633 device.
7. The OLT reports an event to the U2000, indicating that the MA5633 device has been online.
8. The MA5633 device initiates a registration request to the OLT.
9. The OLT reports an event to the U2000, indicating that the OLT has automatically discovered the MA5633 device.
10. The software commissioning personnel run the **frame bind** command on the OLT to bind the SN of the MA5633 device to the extended frame. In this manner, the extended frame is added to the OLT.
11. **Optional:** The OLT automatically upgrades the MA5633 software if the MA5633 software version is different that the OLT software version.
12. The OLT reports an event to the U2000, indicating that the MA5633 device is functioning properly.
13. **Optional:** If an RF line initialization profile has been configured on the OLT, the OLT automatically issues RF line configurations to the MA5633 device according to the profile after the OLT adds the MA5633 device.
14. The software commissioning personnel initiate a DHCP emulation test through the U2000 to verify the data configuration. Then, the data configuration is complete.

NOTE

If the software commissioning personnel have obtained the SN of the MA5633, they can perform the 3 step in **Remote GPON Extended Frame Deployment in Offline Mode** to bind the SN of the MA5633 device to the extended frame. Then the 10 step is omitted. Other steps remain unchanged.

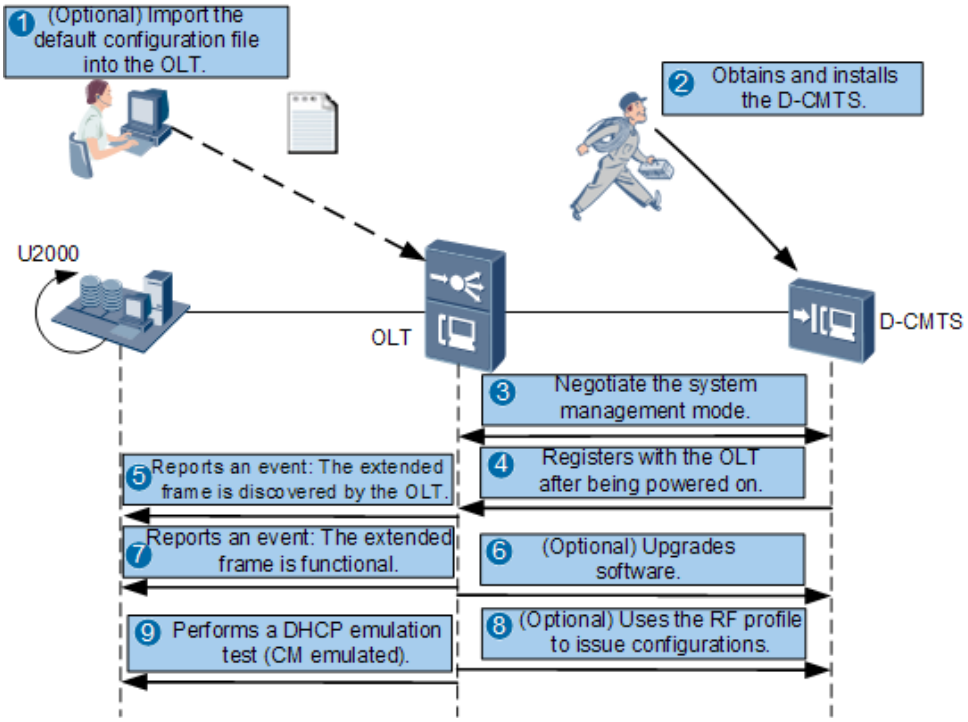
33.4.4 Centralized Management for GE Extended Frames

Automatic Discovery for GE Extended Frames

The automatic discovery for GE extended frames enables an optical line terminal (OLT) to automatically acknowledge and add a GE extended frame. This simplifies the flows for registering and adding an extended frame. This function applies to the scenario where a great number of distributed cable modem termination systems (MA5633s) need to be deployed.

Figure 33-9 shows the process of deploying automatic discovery for GE extended frames.

Figure 33-9 Process of deploying automatic discovery for GE extended frames



1. **Optional:** The software commissioning personnel import the default configuration file, which includes RF parameter settings of the MA5633, into the OLT.

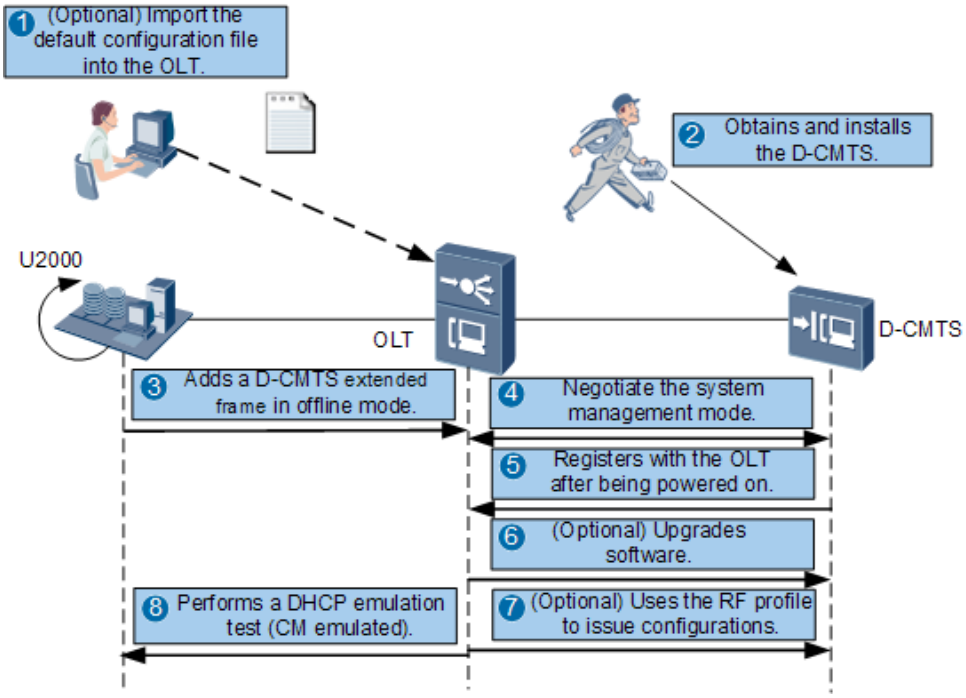
   Run the **display network-role** command on the OLT to check the working mode of the ETHB cascading board. If the working mode is not **extend**, run the **network-role** command to change the working mode to **extend**.
2. The hardware installation personnel obtain the data planning table, fetch the MA5633 device according to the MAC address in the data planning table, and install the MA5633 device on site. The hardware installation personnel power on the MA5633 device and make optical paths available.
3. The MA5633 device and the OLT negotiate the system management mode. For details, see **System Management Mode Autonegotiation** in 33.4.2 Basic Concepts.
4. The MA5633 device initiates a registration request to the OLT and the OLT adds the MA5633 device as an extended frame.
5. The OLT reports an event to the U2000, indicating that the OLT has automatically discovered the MA5633 device.
6. **Optional:** The MA5633 device automatically upgrades software.
7. The OLT reports an event to the U2000, indicating that the MA5633 device is functioning properly.
8. **Optional:** If a radio frequency (RF) line initialization profile has been configured on the OLT, the OLT automatically issues RF line configurations to the MA5633 device according to the profile after the OLT adds the MA5633 device.
9. The software commissioning personnel initiate a Dynamic Host Configuration Protocol (DHCP) emulation test through the U2000 to verify the data configuration. Then, the data configuration is complete.

GE Extended Frame Deployment in Offline Mode

Deploying GE extended frames in offline mode applies to the scenario where carriers plan MA5633 extended frame IDs in a unified manner and therefore need to deploy the MA5633s one by one. This facilitates device management.

Figure 33-10 shows the process of deploying a GE extended frame in offline mode.

Figure 33-10 Process of deploying a GE extended frame in offline mode



1. **Optional:** The software commissioning personnel import the default configuration file, which includes RF parameter settings of the MA5633, into the OLT.
2. The hardware installation personnel obtain the data planning table, fetch the MA5633 device according to the MAC address in the data planning table, and install the MA5633 device on site.
3. The software commissioning personnel run the **network-role** command on the OLT to configure the working mode of the ETHB cascading board to **extend** and run the **frame add** command to add the MA5633 device in offline mode.
4. The hardware installation personnel power on the MA5633 device and make optical paths available. The MA5633 device and the OLT negotiate the system management mode. For details, see **System Management Mode Autonegotiation** in 33.4.2 Basic Concepts.
5. The MA5633 device initiates a registration request to the OLT and the OLT adds the MA5633 device as an extended frame.
6. **Optional:** The MA5633 device automatically upgrades software.
7. **Optional:** If an RF line initialization profile has been configured on the OLT, the OLT automatically issues RF line configurations to the MA5633 device according to the profile after the OLT adds the MA5633 device.

8. The software commissioning personnel initiate a DHCP emulation test through the U2000 to verify the data configuration. Then, the data configuration is complete.

33.5 PacketCable

PacketCable is formulated by CableLabs with the goal of providing a standard architecture to implement VoIP and IP-based multimedia services over cable networks.

Introduction

In addition to providing video and data services, multiservice operators (MSOs) expect to transmit VoIP services over existing cable networks to fully leverage the service growth potential of the networks. This requires a set of standards for VoIP service implementation.

PacketCable defines a component-based architecture and a set of interfaces supporting IP-based communication technologies, such as the Session Initiation Protocol (SIP), to provide integrated real-time multimedia services, such as voice, video, data, and mobile services, over cable TV (CATV) networks.

PacketCable consists of PacketCable 1.0, PacketCable 1.5, PacketCable Multimedia, and PacketCable 2.0.

PacketCable 1.0	Used to provide voice service through embedded media terminal adapters (EMTAs).
PacketCable 1.5	Developed based on PacketCable 1.0 and uses SIP to manage sessions on PacketCable networks. **NOTE** PacketCable 1.0 and PacketCable 1.5 use the same architecture and are represented as PacketCable 1.x in the rest of the document.
PacketCable Multimedia	Provides a billing architecture and quality of service (QoS) guarantees that are independent of services, unlike PacketCable 1.x. Therefore, PacketCable Multimedia can be used to provide QoS guarantees for various applications, such as gaming, audio, and video services.
PacketCable 2.0	Developed for SIP-based clients and upper-layer service platforms to provide more service types. **NOTE** PacketCable 2.0 does not apply to access layer devices and therefore is not described in the remainder of this document.

33.5.1 PacketCable 1.x

This section describes the architecture, interfaces, and implementation principles of PacketCable 1.x.

Introduction

PacketCable 1.x supports voice services over cable networks but does not support multimedia services, such as video services.

PacketCable 1.x Architecture

Figure 33-11 shows the PacketCable 1.x architecture.

Figure 33-11 PacketCable 1.x architecture

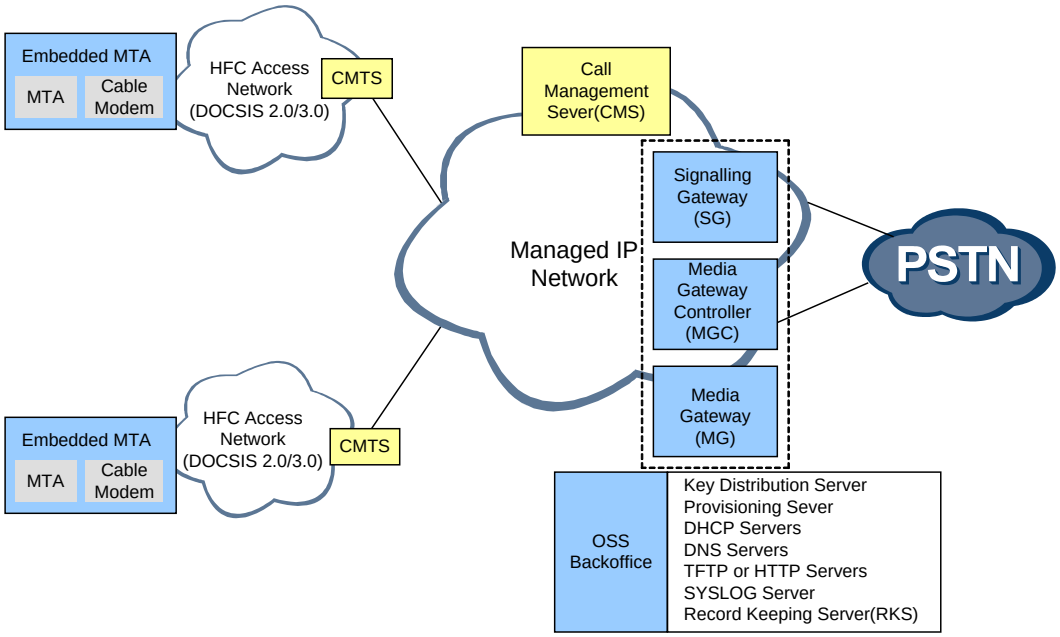


Table 33-10 describes the components in the preceding figure and their functions.

Table 33-10 PacketCable 1.x components and functions

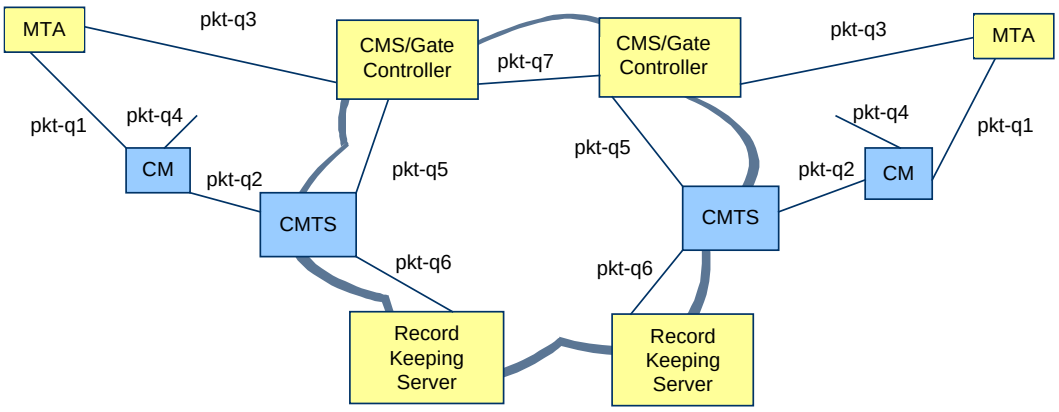
Component	Function
Media terminal adapter (MTA)	Provides a user port on the user side for a customer premises equipment (CPE) device, such as a phone set, and a signaling interface on the network side for a call control unit. The MTA connects to a cable network unit through the hybrid fiber coaxial (HFC) network. PacketCable classifies MTAs as EMTAs and standalone MTAs (SMTAs). • An EMTA integrates cable modem (CM) and MTA functions. • An SMTA supports only MTA functionalities. **NOTE** The PacketCable protocol defines only the interfaces on an EMTA network. The MA5600T/MA5603T/MA5608T supports only EMTAs. MTAs support the Network-Based Call Signaling (NCS) protocol. **NOTE** NCS, extended based on the Media Gateway Control Protocol version 1.0 (MGCPv1), is used between EMTAs and CMSs to meet the requirements of cable applications and QoS.
Call management server	Functions as a switch on PSTN networks or a softswitch on NGN networks. It exchanges signaling with MTAs through the NCS protocol. In this manner, multimedia connections are set up between the MTAs and other

Component	Function
(CMS)	devices on the network.
Cable modem termination system (CMTS)	Connects a data network to an HFC network. It forwards network data, processes protocols, and modulates and demodulates radio frequency (RF) signals. The CMTS provides QoS guarantees required by CMs and allocates upstream bandwidth and service resources to the CMs based on CM requests and network QoS policies.
Cable modem (CM)	Connects a CPE to an HFC network. It transmits data over cable networks through the data over cable service interface specification (DOCSIS) protocol.
Media gateway controller (MGC), media gateway (MG), and signaling gateway (SG)	Constitute a gateway that connects a cable network to a PSTN network. The functions of the three components are the same as their functions on traditional PSTN networks. Therefore, this document does not describe the three components.

PacketCable 1.x Interfaces

The transmission of voice and multimedia signals over cable networks requires high QoS guarantees. Figure 33-11 shows the QoS interfaces used by PacketCable 1.x.

Figure 33-12 PacketCable 1.x interfaces



The following section describes the new concepts involved in the preceding figure and their functions.

Gate controller (GC)

A GC is integrated into a CMS and works as a functional component of the CMS. The GC controls QoS and exchanges data with a CMTS to determine whether to accept a QoS request initiated from an MTA.

Record keeping server (RKS)

An RKS only receives messages from other components. It assembles event messages as a set or classifies them into call detail records (CDRs) for server processing, such as error correction or billing.

Table 33-11 briefly describes the PacketCable 1.x interfaces and their functions. For more information, see *PKT-SP-DQOS1.5-I04-090624* released by CableLabs.

Table 33-11 PacketCable 1.x interfaces and functions

Interface	Device Connected by the Interface	Function
pkt-q1	MTA-CM	Used for implementing the following functions between CMs and EMTAs: traffic control on service flows, data packet synchronization and transmission, and QoS guarantees. **NOTE** MTAs are used to refer only to EMTAs in the remainder of this document.
pkt-q2	CM-CMTS	A DOCSIS radio frequency interference (RFI) QoS interface, used for controlling, scheduling, and transmitting packets. The control function can be initiated from either CMs or CMTSs. Only the CMTSs have the capability of determining the control policy.
pkt-q3	MTA-GC or MTA-CMS	Used for transmitting NCS signaling.
pkt-q4	CM-Provisioning server	Not used in PacketCable 1.x.
pkt-q5	GC-CMTS	Used for managing dynamic gates in media stream sessions through COPS. For details about the COPS protocol, see 33.5.3 COPS.
pkt-q6	CMTS-RKS	Used by CMTSs to notify RKSs of session changes during authorization and usage.
pkt-q7	CMS-CMS	Used for managing sessions and coordinating resources between CMTSs.

Principles

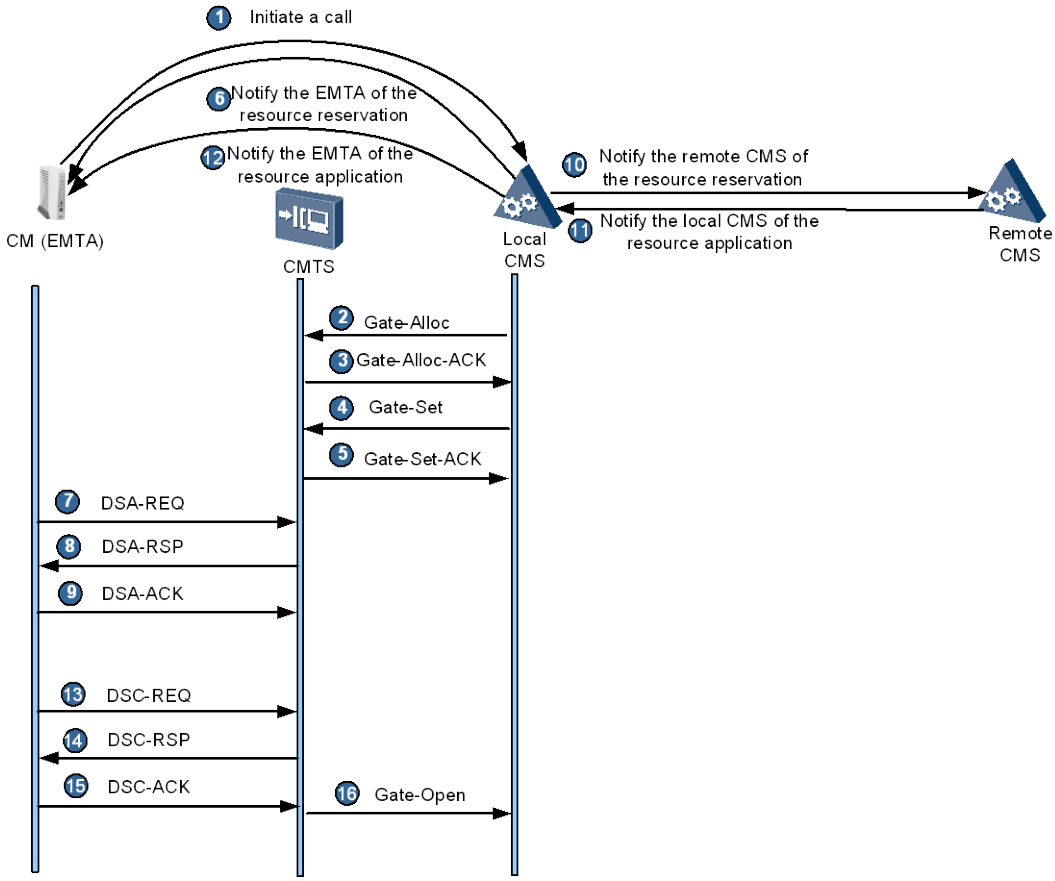
This section describes the PacketCable 1.x principles for setting up and terminating a voice session.

NOTE
In the following figures, the CMTS refers to an MA5633 in standalone NE mode or an optical line terminal (OLT) in centralized management mode.

Setting Up a Voice Session

Figure 33-13 shows the process of setting up a voice session.

Figure 33-13 Process of setting up a voice session



In the preceding figure, the process involves three phases.

1. Resource authorization phase (steps 1 through 5)
   a. The EMTA initiates a call request to the local CMS.
   b. The local CMS sends a Gate-Alloc message to the MA5600T/MA5603T/MA5608T to request a gate ID. The Gate-Alloc message carries the number of gates (indicated by the **Activity-Count** value) supported by the EMTA.
   c. The MA5600T/MA5603T/MA5608T receives the Gate-Alloc message and checks whether the number of gates that have been created on the MA5600T/MA5603T/MA5608T reaches the gate threshold configured on the MA5600T/MA5603T/MA5608T, and whether the number of gates that have been created on the EMTA reaches the gate threshold configured on the EMTA.
      - If neither threshold has been reached, the MA5600T/MA5603T/MA5608T allocates a gate ID to the EMTA, sends a Gate-Alloc-ACK message to the local CMS, and sets the status of the gate state machine to **Allocated**.
      - If either threshold has been reached, the MA5600T/MA5603T/MA5608T cannot allocate a gate ID to the EMTA and sends a Gate-Alloc-Err message to the local CMS. Then, the session fails to be set up.

d. The local CMS receives the Gate-Alloc-ACK message, configures QoS parameters required by the session, and sends a Gate-Set message to the MA5600T/MA5603T/MA5608T for resource authorization.

e. The MA5600T/MA5603T/MA5608T receives the Gate-Set message and creates a gate for the EMTA. Then, the MA5600T/MA5603T/MA5608T saves the data in a spreadsheet, sends a Gate-Set-ACK message to the local CMS, and sets the status of the gate state machine to **Authorized**.

2. Resource reservation phase (steps 6 through 11)

a. The local CMS notifies the EMTA of the resource reservation for the session and issues the traffic classification and QoS parameters as well as the gate ID to the EMTA.

b. The EMTA sends a DSA-REQ message to the MA5600T/MA5603T/MA5608T to request resource reservation. The DSA-REQ message carries the traffic classification and QoS parameters, gate ID, and service flow status (**Admitted**).

c. The MA5600T/MA5603T/MA5608T receives the DSA-REQ message, obtains the gate ID from the message, and checks whether the resources requested by the EMTA are authorized by the local CMS and whether the remaining resources meet the EMTA requirements.

- If all of the requirements are met, the MA5600T/MA5603T/MA5608T creates a service flow for the EMTA, sends a DSA-RSP message to the EMTA, and sets the status of the gate state machine to **Reserved**.
- If any of the requirements is not met, the MA5600T/MA5603T/MA5608T sends a DSA-RSP message to reject the EMTA request.

d. The EMTA receives the DSA-RSP message and sends a DSA-ACK message to the MA5600T/MA5603T/MA5608T.

e. The local CMS notifies the remote CMS of the resource authorization and reservation.

f. The remote CMS reserves the resources and notifies the local CMS of the resource allocation to the EMTA.

3. Resource allocation phase (steps 12 through 16)

a. The local CMS notifies the EMTA of the resource allocation request.

b. The EMTA sends a DSC-REQ message to the MA5600T/MA5603T/MA5608T to request resource allocation. The DSC-REQ message carries the traffic classification and QoS parameters, gate ID, and service flow status (**Active**).

c. The MA5600T/MA5603T/MA5608T receives the DSC-REQ message and obtains the gate ID from the message. Then, the MA5600T/MA5603T/MA5608T checks whether the resources requested by the EMTA are less than or equal to the reserved resources.

- If the reserved resources meet the requirements, the MA5600T/MA5603T/MA5608T changes the status of the service flow to **Active**, sends a DSC-RSP message to the EMTA, and sets the status of the gate state machine to **Committed**.
- If the reserved resources do not meet the requirements, the MA5600T/MA5603T/MA5608T sends a DSC-RSP message to the EMTA to reject the request, and the session fails to be set up.

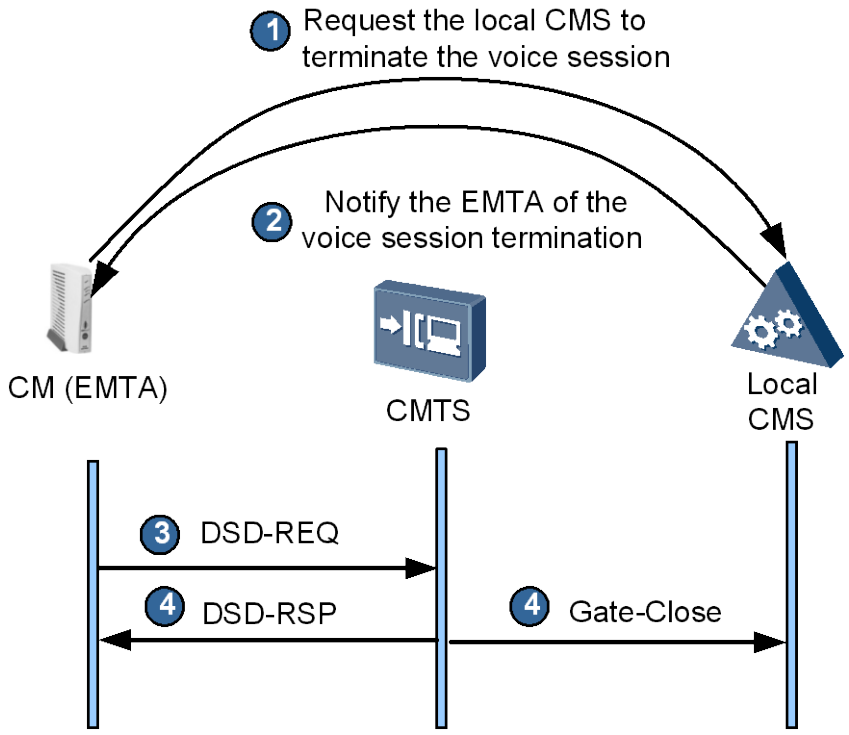
d. The EMTA receives the DSC-RSP message and sends a DSC-ACK message to the MA5600T/MA5603T/MA5608T.

e. The MA5600T/MA5603T/MA5608T sends a Gate-Open message to the local CMS, notifying the local CMS of successful resource allocation. The voice session is set up.

Terminating a Voice Session

Figure 33-14 shows the process of terminating a voice session.

Figure 33-14 Process of terminating a voice session



The process is as follows:

1. The user hangs up the phone. Then, the EMTA requests the local CMS to terminate the voice session.
2. The local CMS sends a message to the EMTA to notify the EMTA of the voice session termination and starts the timer.
3. The EMTA sends a DSD-REQ message to the MA5600T/MA5603T/MA5608T to request a service flow deletion.
4. The MA5600T/MA5603T/MA5608T receives the DSD-REQ message, deletes the gate service flow and other resources, sends a DSD-RSP message to the EMTA, and sends a Gate-Close message to the local CMS.
5. The local CMS receives the Gate-Close message and deletes the voice session.

NOTE

If the local CMS does not receive the Gate-Close message before its timer expires, the local CMS sends a Gate-Delete message to the MA5600T/MA5603T/MA5608T. Then, the MA5600T/MA5603T/MA5608T deletes the gate service flow and other resources and sends a Gate-Delete-ACK message to the local CMS.

33.5.2 PacketCable Multimedia

This section describes the architecture, interfaces, and implementation principles of PacketCable Multimedia.

Introduction

PacketCable Multimedia defines a QoS architecture independent of services and provides QoS guarantees for various applications, such as gaming, audio, and video services. PacketCable Multimedia provides enhanced QoS for multimedia services than PacketCable 1.x and implements QoS protocol conversion using application managers (AMs) and policy servers (PSs).

PacketCable Multimedia Architecture

Figure 33-15 shows the PacketCable Multimedia architecture.

Figure 33-15 PacketCable Multimedia architecture

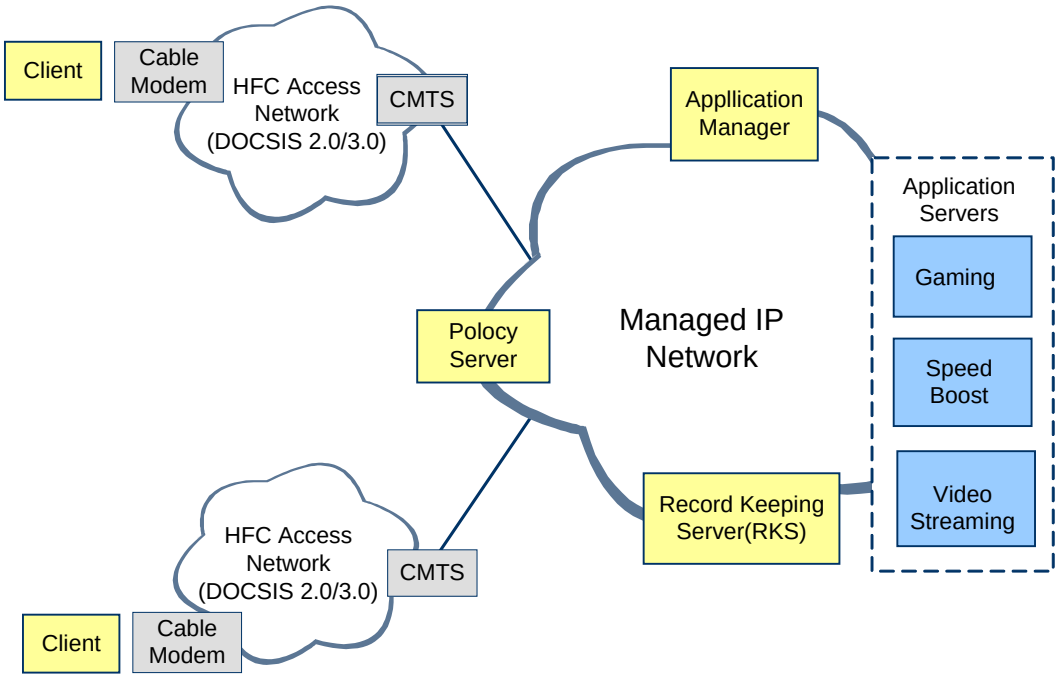


Table 33-12 describes the components in the preceding figure and their functions.

Table 33-12 PacketCable Multimedia components and functions

Component	Function
Client	Initiates or terminates a call. The client can be any terminal that can initiate an audio or video request, such as a phone set or PC.
AM	Parses a voice session and determines the resources required for the session. The AM works with a PS to implement the policy and charging rules

Component	Function
	function (PCRF).
PS	Determines the rights of the call-initiating user and session resources.
CMTS	Connects a data network to an HFC network. It forwards network data, processes protocols, and modulates and demodulates radio frequency (RF) signals. The CMTS provides QoS guarantees required by CMs and allocates upstream bandwidth and service resources to the CMs based on CM requests and network QoS policies.
CM	Connects a CPE to an HFC network. It transmits data over cable networks through the data over cable service interface specification (DOCSIS) protocol.
RKS	An RKS only receives messages from other components. It assembles event messages as a set or classifies them into call detail records (CDRs) for server processing, such as error correction or billing.

PacketCable Multimedia Interfaces

The transmission of voice and multimedia signals over cable networks requires high QoS guarantees. Figure 33-16 shows the QoS interfaces used by PacketCable Multimedia.

Figure 33-16 PacketCable Multimedia interfaces

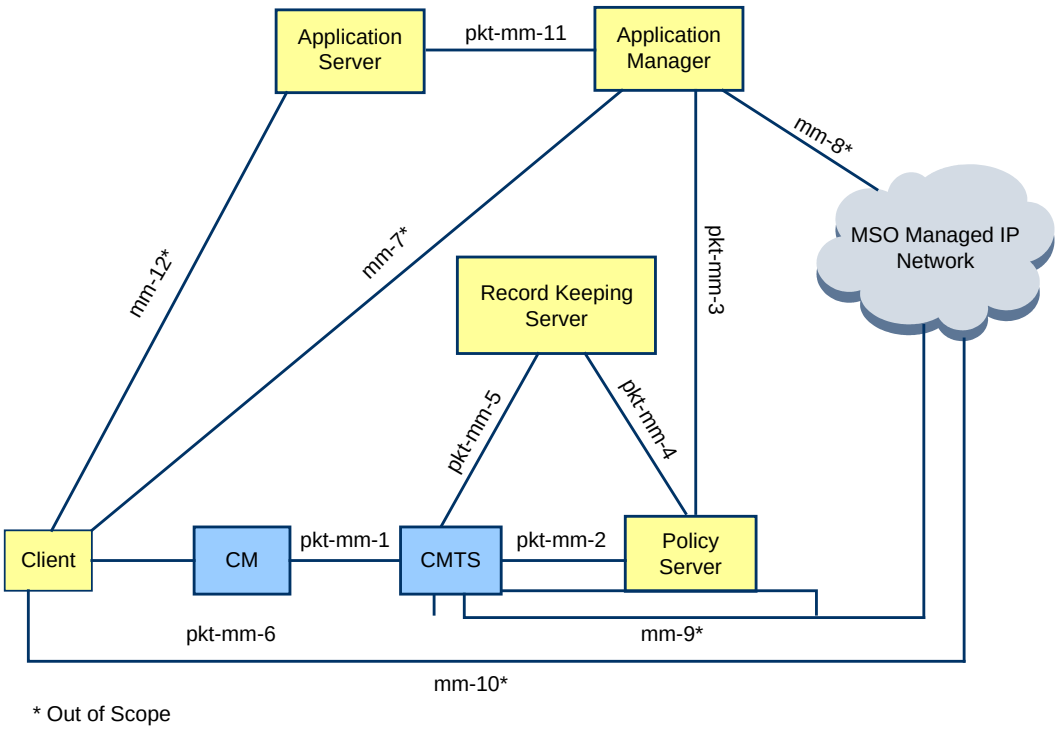


Table 33-13 describes the PacketCable Multimedia interfaces and their functions.

NOTE

In the preceding figure, the interfaces marked with an asterisk (*) have not been defined and therefore are not included in the following table.

Table 33-13 PacketCable Multimedia interfaces and functions

Interface	Device Connected by the Interface	Function
pkt-mm-1	CMTS-CM	Uses DOCSIS-compliant DSx signaling to create, modify, or delete service flows that meet DOCSIS QoS request requirements between the CM and the MA5600T/MA5603T/MA5608T.
pkt-mm-2	PS-CMTS	Uses COPS for sending a PS policy to the MA5600T/MA5603T/MA5608T, or for sending a message to a PS to notify the PS of MA5600T/MA5603T/MA5608T resource changes. For details about the COPS protocol, see 33.5.3 COPS.
pkt-mm-3	AM-PS	Uses COPS for sending a policy setting request to a PS from an AM, or for sending a message to an AM to notify the AM of QoS resource status changes. For details about the COPS protocol, see 33.5.3 COPS.
pkt-mm-4	PS-RKS	Used by a PS to send voice transactions to an RKS.
pkt-mm-5	CMTS-RKS	Used by the MA5600T/MA5603T/MA5608T to send voice transactions to an RKS.
pkt-mm-6	Client-CMTS	Used by a client to send QoS resource request and management packets to the MA5600T/MA5603T/MA5608T. The client can obtain the QoS resources after being authorized.
pkt-mm-11	AS-AM	Used for sending the policy setting messages exchanged between a proxy-call session control function (P-CSCF) and an AM, or for sending a message to an R-CSCF to notify the P-CSCF of QoS resource status changes. **NOTE** The P-CSCF, as defined in PacketCable 2.0, manages user access and QoS resources.

Principles

This section describes the PacketCable Multimedia principles for setting up and terminating a voice session.

The CMTS supports two types of PacketCable Multimedia-compliant clients.

- Client 1: includes PCs, gaming consoles, and Session Initiation Protocol (SIP) terminals. Clients of this type cannot instruct CMs to initiate DSx signaling for creating service flows.

- Client 2: refers to the terminals that are similar to PacketCable 1.x terminals and support QoS functions. Clients of this type can instruct CMs to initiate DSx signaling for creating service flows.

NOTE

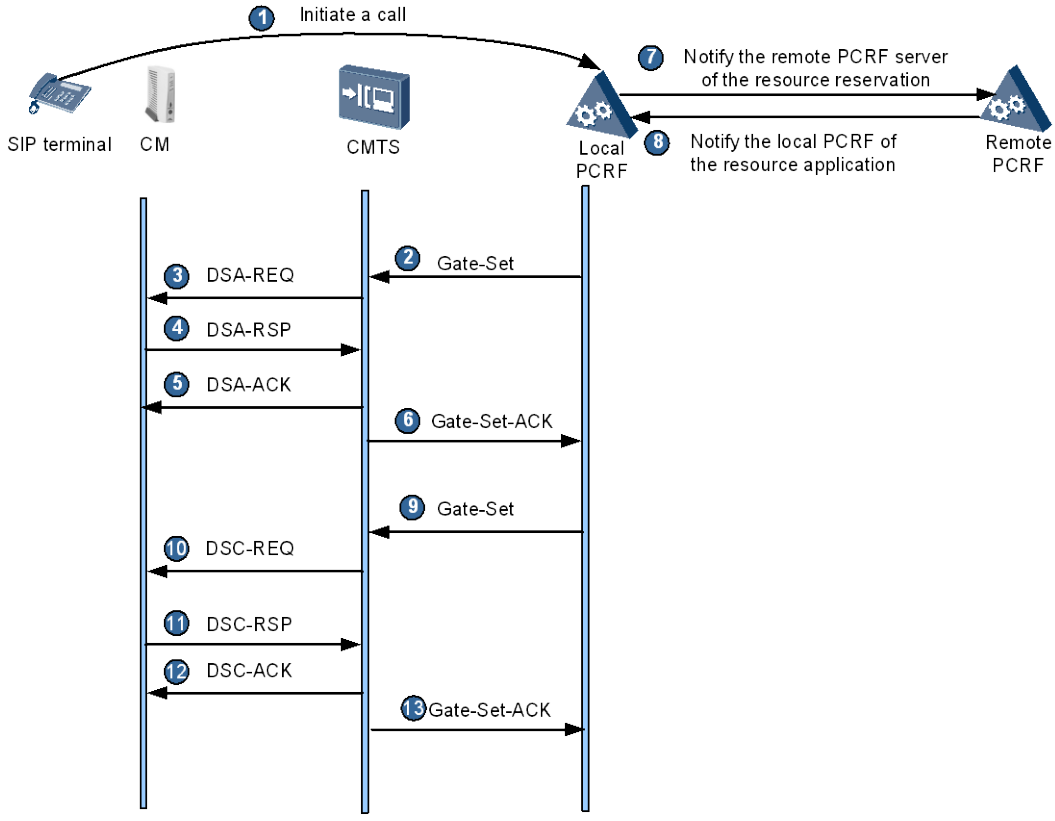
- For client 2, the voice session setup and termination processes in PacketCable Multimedia are the same as those in PacketCable 1.x and therefore the description is not provided in this document. The following section only describes the voice session setup and termination processes in PacketCable Multimedia for client 1.
- In the following figures, the CMTS refers to an MA5633 in standalone NE mode or an OLT in centralized management mode.

Setting Up a Voice Session

A voice session can be set up by creating a service flow in one or two steps.

Figure 33-17 shows the process of setting up a voice session by creating a service flow in two steps.

Figure 33-17 Process of setting up a voice session by creating a service flow in two steps



In the preceding figure, the process involves two phases.

1. Resource authorization and reservation phase (steps 1 through 8)
   a. The SIP terminal initiates a call request to the local PCRF server.
   b. The local PCRF server configures the QoS parameters required by the session and sends a Gate-Set message to the CMTS to authorize and reserve resources for the CMTS.

c. The CMTS receives the Gate-Set message and creates a gate for the SIP terminal. Then, the CMTS saves the data in a spreadsheet, creates a service flow, sends a DSA-REQ message to the CM to create a dynamic service flow, and sets the status of the gate state machine to **Admitted**.

d. The CM sends a DSA-RSP message to the CMTS.

e. The CMTS receives the DSA-RSP message and sends a DSA-ACK message to the CM.

f. The CMTS sends a Gate-Set-Ack message to the local PCRF server.

g. The local PCRF server notifies the remote PCRF server of the resource reservation.

h. The remote PCRF server reserves the resources and notifies the local PCRF of the resource allocation to the SIP terminal. In addition, the remote SIP terminal receives the call.

2. Resource allocation phase (steps 9 through 13)

a. The local PCRF server sends a Gate-Set message to the CMTS to notify the CMTS of the resource allocation request.

b. The CMTS sends a DSC-REQ message to the CM to change the status of the dynamic service flow to **Active**.

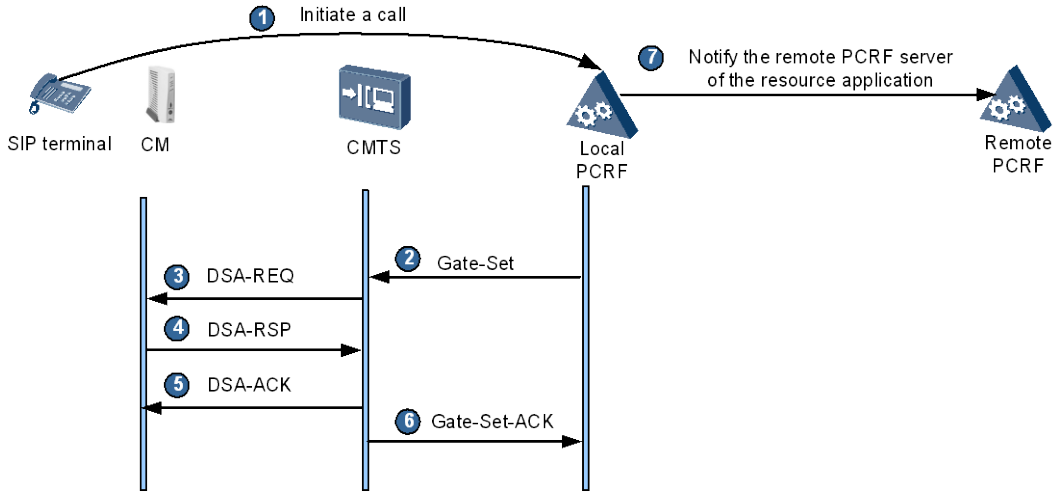
c. The CM sends a DSC-RSP message to the CMTS.

d. The CMTS receives the DSC-RSP message and sends a DSC-ACK message to the CM.

e. The CMTS changes the status of the service flow to **Active** and sends a Gate-Set-Ack message to the local PCRF server to notify the local PCRF server of the successful resource allocation. Then, the voice session is set up.

Figure 33-18 shows the process of setting up a voice session by creating a service flow in one step.

Figure 33-18 Process of setting up a voice session by creating a service flow in one step



In the preceding figure, the process is as follows:

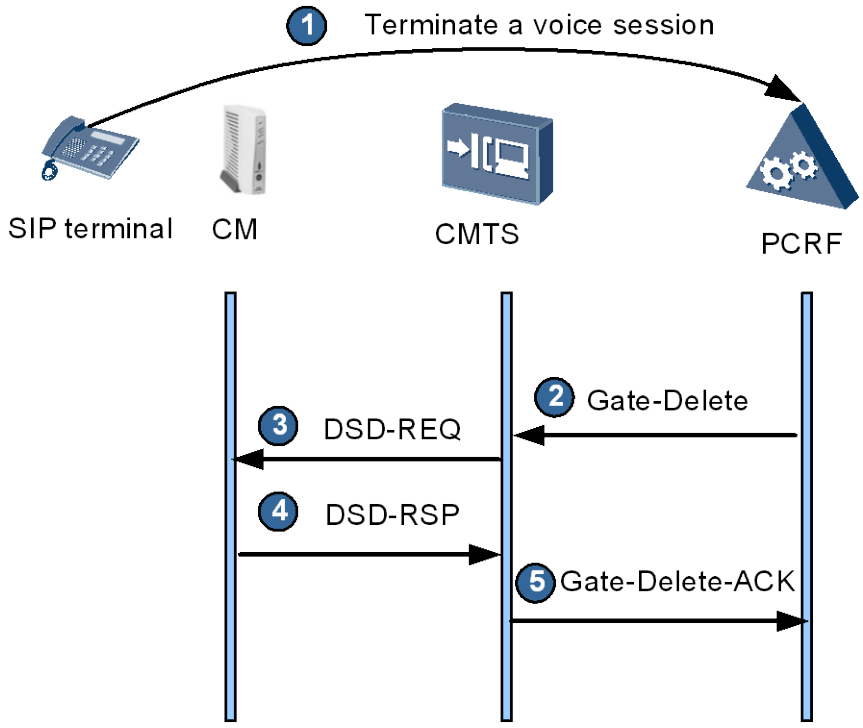
1. The SIP terminal initiates a call request to the local PCRF server.

2. The local PCRF server configures the QoS parameters required by the session and sends a Gate-Set message to the CMTS to authorize and reserve resources for the CMTS and allocate the resources to the CMTS.
3. The CMTS receives the Gate-Set message, initializes the status of the gate state machine to **Authorized**, and creates a gate for the SIP terminal. Then, the CMTS saves the data in a spreadsheet, creates a service flow, and sends a DSA-REQ message to the CM to create a dynamic service flow.
4. The CM sends a DSA-RSP message to the CMTS.
5. The CMTS receives the DSA-RSP message and sends a DSA-ACK message to the CM.
6. The CMTS changes the status of the service flow to **Active** and the status of the gate state machine to **Committed** and sends a Gate-Set-Ack message to the local PCRF server to notify the local PCRF server of the successful resource allocation.
7. The local PCRF server notifies the remote PCRF server of the resource allocation.
8. The remote PCRF server allocates the resources and the remote SIP terminal receives the call. Then, the voice session is set up.

Terminating a Voice Session

Figure 33-19 shows the process of terminating a voice session.

Figure 33-19 Process of terminating a voice session



The process is as follows:

1. The user hangs up the phone. Then, the SIP terminal requests the PCRF server to terminate the voice session.
2. The PCRF server sends a Gate-Delete message to the CMTS to notify the CMTS of the voice session termination.

3. The CMTS sends a DSD-REQ message to the CM to request a service flow deletion.
4. The CM receives the DSD-REQ message and sends a DSD-RSP message to the CMTS.
5. The CMTS receives the DSD-RSP message, deletes the gate service flow and other resources and sends a Gate-Delete-Ack message to the PCRF server.
6. The PCRF server receives the Gate-Delete-Ack message and deletes the voice session.

33.5.3 COPS

This section describes the components and implementation principles of the Common Open Policy Service (COPS) protocol. The COPS protocol plays an important role in controlling QoS policies for PacketCable 1.x and PacketCable Multimedia.

Introduction

COPS transmits PacketCable quality policies and applies to PacketCable scenarios.

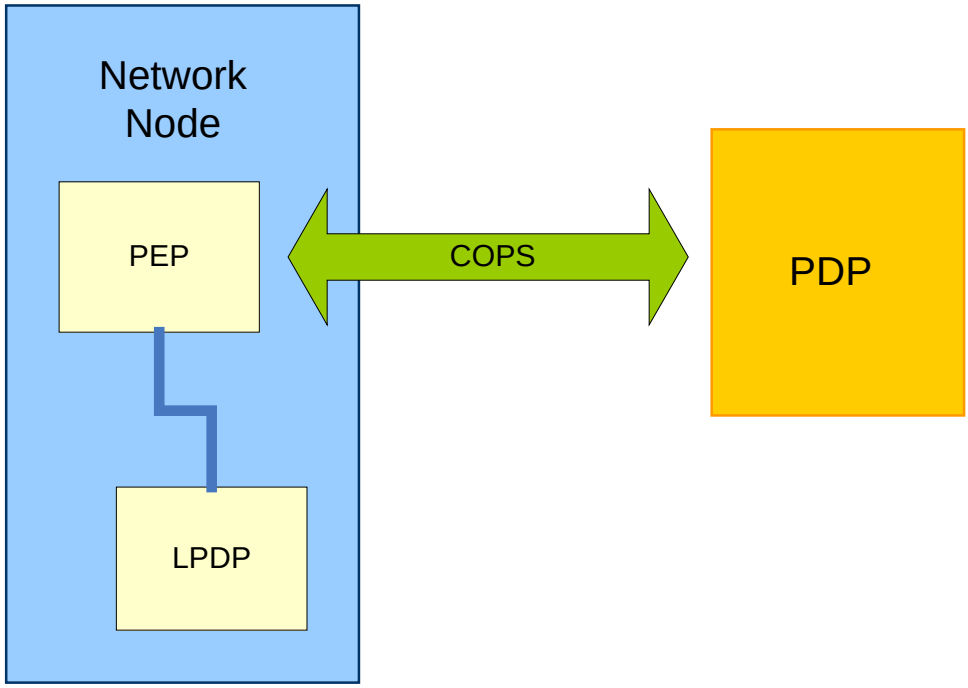
COPS is a simple query and response protocol used between a policy decision point (PDP) and policy enforcement points (PEPs) for exchanging policy information. COPS is used to manage, configure, and implement policies. COPS uses a client/server architecture and runs over TCP to ensure reliable data exchange.

COPS involves three logical entities: PDPs, PEPs, and local policy decision points (LPDPs).

- PDP: controls PEPs. A PDP processes policy information and network resources, determines policies, and sends the policies to PEPs.
- PEP: receives and implements the policies issued from a PDP. PEPs implement the policies when data is transmitted through them.
- LPDP: backs up the policies issued from a PDP. When a PDP disconnects from PEPs, an LPDP uses the backed up PDP polices to manage PEPs.

Figure 33-20 shows the basic COPS model.

Figure 33-20 Basic COPS model



A PDP and a PEP can be regarded as a server and a client, respectively. The PEP sends a configuration, update, or deletion request to the PDP. After receiving the request, the PDP replies to the PEP with a policy. Then, the PEP implements the policy.

COPS uses TCP for data transmission. A PEP initiates a TCP connection to a PDP and periodically sends a Keep_Alive message to the PDP to check the connection validity.

NOTE
In the Huawei PacketCable solutions, the PDP sends unsolicited policies and initiates TCP connections.

Principles

This section describes the principles for creating and deleting a COPS connection.

The differences between creating and deleting a COPS connection in PacketCable Multimedia and in PacketCable 1.x are as follows:

- The TCP listening port number is 3918 in PacketCable Multimedia and 2126 in PacketCable 1.x.
- PSs function as PDPs in PacketCable Multimedia while CMSs function as PDPs in PacketCable 1.x.

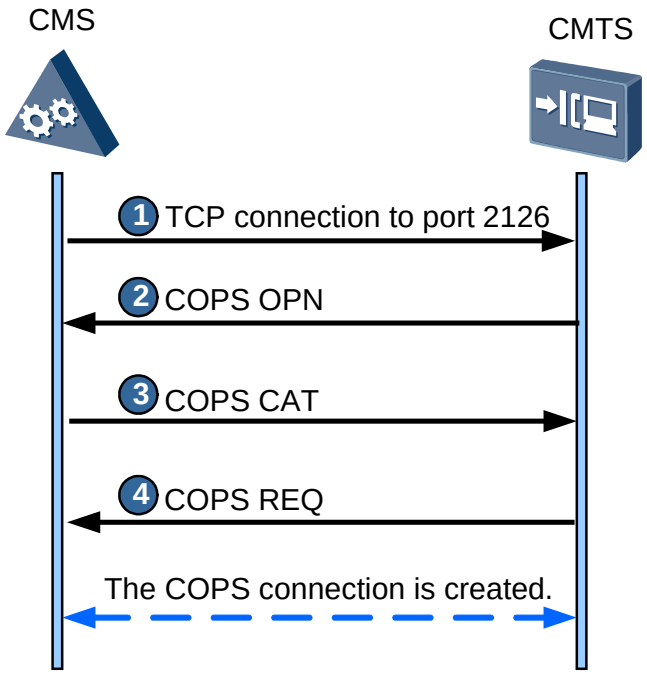
The following section only describes the principles of creating and deleting a COPS connection in PacketCable 1.x.

NOTE
In the following figures, the CMTS refers to an MA5633 in standalone NE mode or an OLT in centralized management mode.

Creating a COPS Connection

Figure 33-21 shows the process of creating a COPS connection.

Figure 33-21 Process of creating a COPS connection



In the preceding figure, the CMS is a PDP and the MA5600T/MA5603T/MA5608T is a PEP. The process is as follows:

1. The MA5600T/MA5603T/MA5608T creates a socket and uses TCP port 2126 on Layer 3 to listen for and receive TCP connection requests from the CMS.
2. The MA5600T/MA5603T/MA5608T sends a Client-Open (COPS OPN) message to the CMS.
3. The CMS receives the Client-Open message and replies with a Client-Accept (COPS CAT) message.
4. The MA5600T/MA5603T/MA5608T sends a COPS Request (COPS REQ) message to the CMS. Then, the COPS connection is created.

After the COPS connection is created, the MA5600T/MA5603T/MA5608T periodically detects the heartbeat for the COPS connection. The process of detecting a heartbeat is as follows:

1. The MA5600T/MA5603T/MA5608T periodically sends a Keep_Alive (KA) message to the CMS according to the timer negotiated between the MA5600T/MA5603T/MA5608T and the CMS.
2. The CMS responds to the message.

Deleting a COPS Connection

A COPS connection can be deleted if any of the following requirements is met:

- The CMS or the MA5600T/MA5603T/MA5608T detects a connection failure through the Keep_Alive message. Both the CMS and the MA5600T/MA5603T/MA5608T can send a Client-Close (CC) message to the peer end to delete the COPS connection.
- Upon receiving an unauthorized Client-Open (OPN) message from a client, the CMS sends a Client-Close (CC) message to the client to delete the COPS connection.
- After refusing a request from a client, the CMS sends a Client-Close (CC) message to the client to delete the COPS connection.
- Upon receiving an unauthorized Client-Accept (CAT) message from a client, the CMS sends a Client-Close (CC) message to the client to delete the COPS connection.

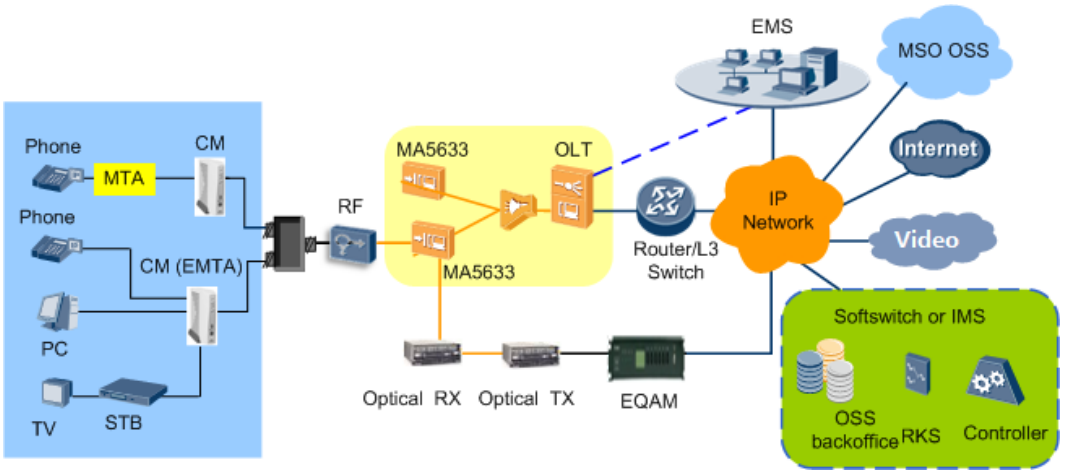
33.5.4 Usage Scenarios

This section describes PacketCable usage scenarios.

PacketCable applies to the following scenarios: centralized management and standalone NE management.

Figure 33-22 shows the networking for centralized management.

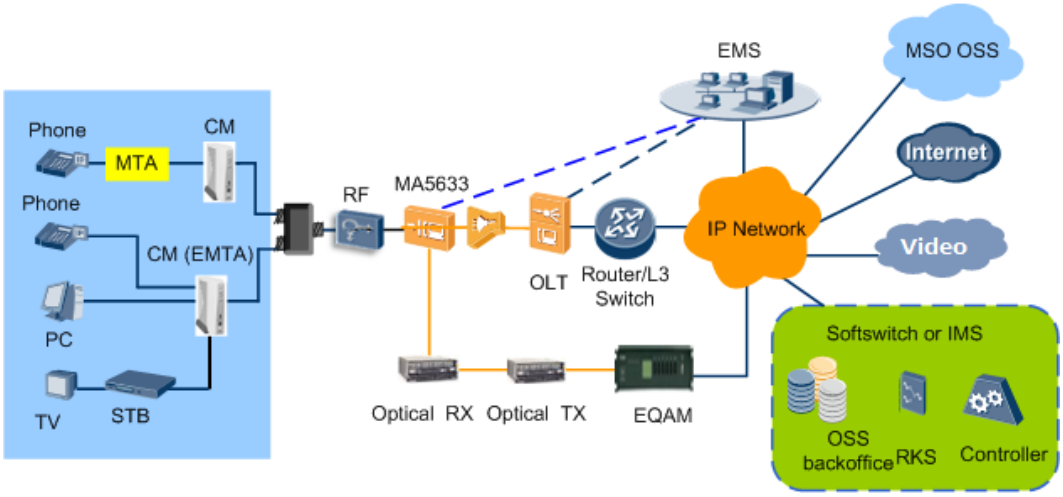
Figure 33-22 Centralized management



In this scenario, the OLT and MA5633s constitute a device and the MA5633s are used as remote extended frames of the OLT. The EMS and core network devices only exchange data with the OLT.

Figure 33-23 shows the networking for standalone NE management.

Figure 33-23 Standalone NE management



In this scenario:

- The MA5633 connects to the OLT through a PON port and the OLT connects to the upper-layer network through a GE port.
- The MA5633 has a separate management IP address and is used as an independent NE. The EMS and core network devices directly exchange data with the MA5633.

33.5.5 Standards and Protocols Compliance

Table 33-14 lists the standards and protocols that the PacketCable feature complies with.

Table 33-14 Standards and protocols that the PacketCable feature complies with

Standard or Protocol	Description
PKT-SP-DQOS1.5-I04-090624	PacketCable™ 1.5 dynamic QoS specifications
PKT-SP-DQOS-C01-071129	PacketCable™ dynamic QoS specifications
PKT-SP-MM-I06-110629	PacketCable™ Multimedia specifications
PKT-SP-MM-WS-I03-091029	PacketCable™ Multimedia specifications (PacketCable Multimedia web service interface specifications)
RFC 2749	Common Open Policy Service (COPS) usage for the Resource Reservation Protocol (RSVP)
RFC 3084	COPS usage for policy provisioning
RFC 2940	Definitions of managed objects for COPS clients
RFC 3483	Policy usage feedback framework for COPS-PR
RFC 3318	Policy information base framework

33.6 Multiple Services in Multiple VLANs

Overview

With the development of hybrid fiber coaxial networks, CMTSs are evolving from traditional CMTSs to coaxial media converters (CMCs) and the deployment location has been moved downwards from distribution hubs to fiber nodes. In this case, the application of the CMTSs on traditional telecom networks poses the following challenges:

- The number of access network layers is changed from two to five. Specifically, the original layers CMTSs and cable modems (CMs) are enlarged to routers, convergence LAN switches, CMCs, CMs, and home gateways (HGWs). The enlarged Layer 2 network promotes higher requirements on the CMCs for processing VLANs.
- Some carriers deploy HGWs on the lower-layer of CMs. The HGWs add different VLAN tags to packets based on service types. This requires that the CMCs support transparent transmission of VLAN tags.

Values

VLAN classification based on service types brings the following business value points for carriers:

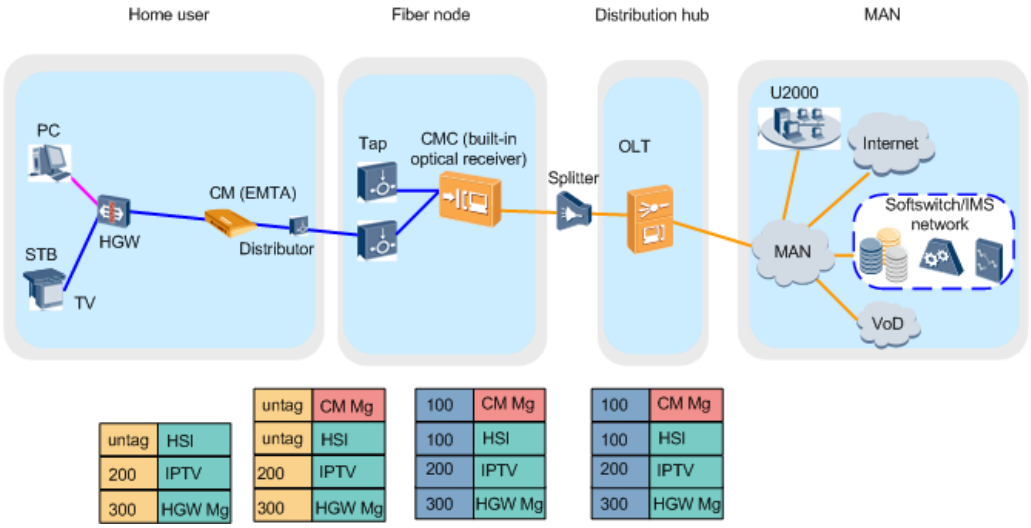
- Services are isolated, which narrows down the broadcast scope.
- Service types can be identified based on VLANs.
- Layer 2 networks can be managed based on VLANs.
- Forwarding paths can be selected on Layer 2 networks based on VLANs.

Typical Network Planning

NOTE

The VLAN tags used in the following section are for reference only.

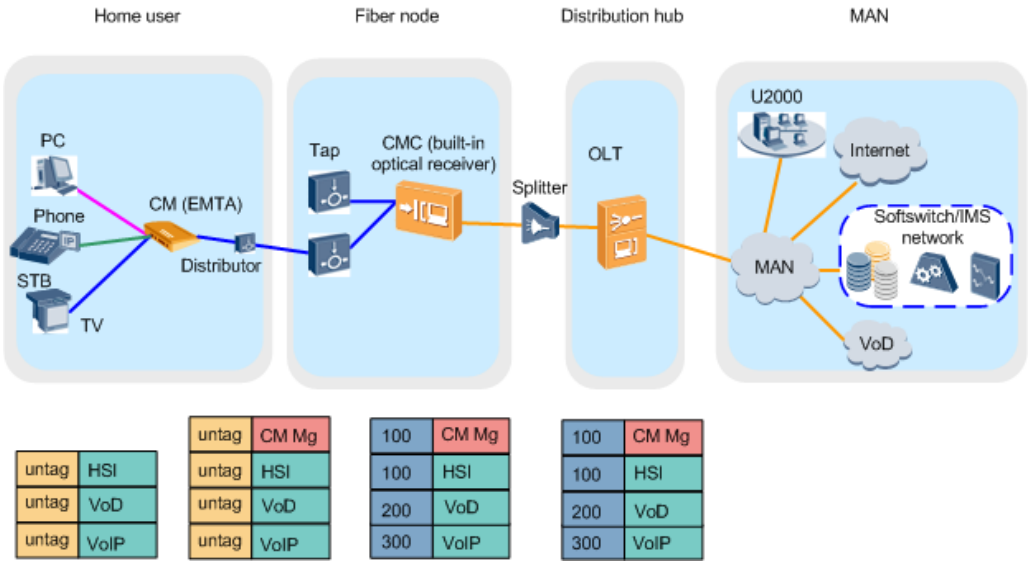
Figure 33-24 VLAN tag transparent transmission (home access)



In the preceding network, IP addresses are allocated and service flows and VLAN tags are processed as follows:

1. The PC, set top box (STB), HGW, and CM use dynamic IP addresses allocated by a Dynamic Host Configuration Protocol (DHCP) server.
2. All packets sent by the PC and STB are untagged. The HGW adds VLAN tag 200 to STB packets, adds VLAN tag 300 to HGW management packets (HGW Mg), and transparently transmits high-speed Internet (HSI) service packets.
3. The CMC adds VLAN tag 100 to the HSI service packets and CM management packets (CM Mg).

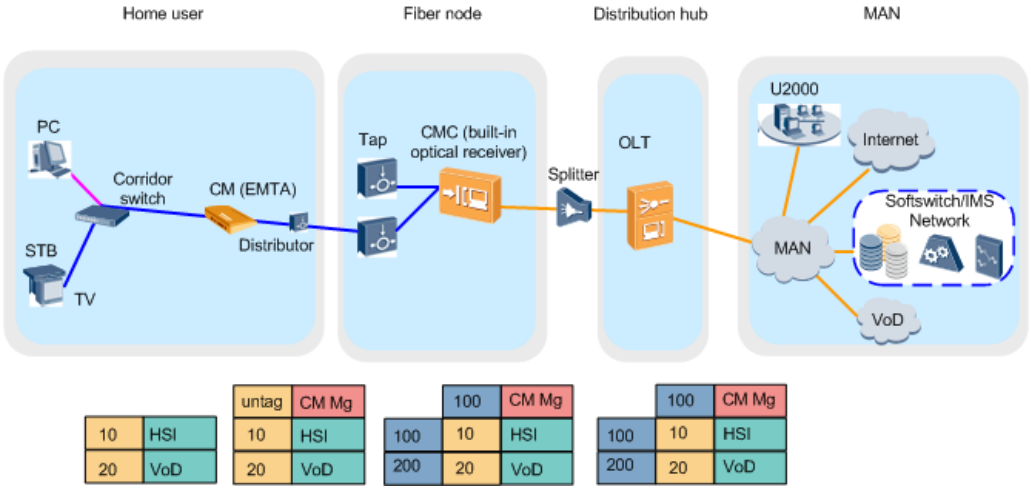
Figure 33-25 Multiple services in multiple VLANs (home access)



In the preceding network, IP addresses are allocated and service flows and VLAN tags are processed as follows:

1. The PC, STB, and CM use dynamic IP addresses allocated by a DHCP server.
2. All packets sent by the PC and STB are untagged.
3. The CMC adds VLAN tag 100 to the HSI service packets and CM management packets (CM Mg), adds VLAN tag 200 to the VoD service packets, and adds VLAN tag 300 to the VoIP service packets.

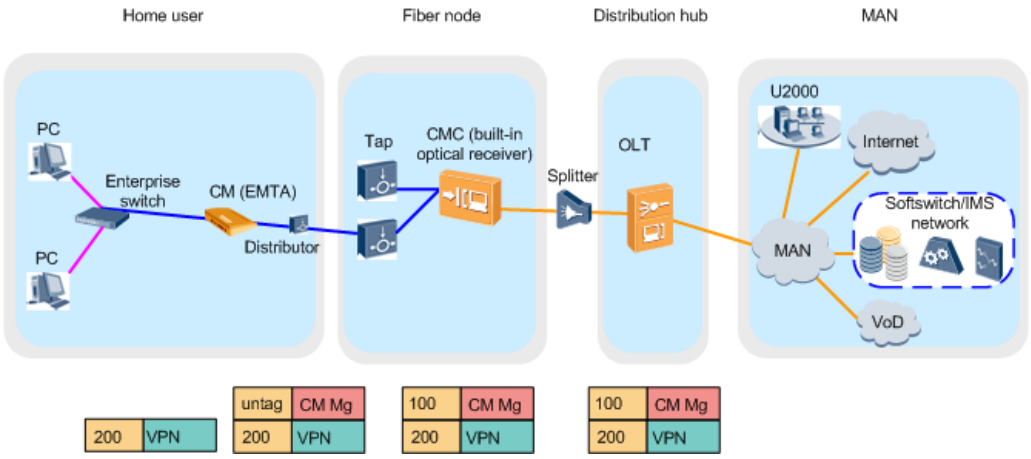
Figure 33-26 VLAN QinQ (home access)



In the preceding network, IP addresses are allocated and service flows and VLAN tags are processed as follows:

1. The CM and STB use dynamic IP addresses allocated by a DHCP server. The PC uses PPPoE dialup.
2. All packets sent by the PC and STB are untagged. The corridor switch adds VLAN tag 10 to the HSI service packets and adds VLAN tag 20 to the VoD service packets.
3. The CMC adds VLAN tag 100 to the HSI service packets and CM management packets (CM Mg) and adds VLAN tag 200 to the VoD service packets.

Figure 33-27 VLAN tag transparent transmission (VPN access)



In the preceding network, IP addresses are allocated and service flows and VLAN tags are processed as follows:

1. The CM uses the dynamic IP address allocated by a DHCP server.
2. The enterprise switch statically configures the IP address of the PC and adds VLAN tag 200 to VPN data packets.
3. The CMC adds VLAN tag 100 to the CM management packets (CM Mg) and transparently transmits the VLAN tag of the VPN data packets.

NOTE

The preceding VLAN planning applies to the CMC, regardless of whether the CMC is deployed in standalone NE or centralized management mode.

VLAN translation policies

Table 33-15 VLAN translation policies supported (home access and VPN access)

VLAN Tag Before Translation	VLAN Tag After Translation	VLAN Translation Policy	Description
Untagged	Service VLAN (S-VLAN)	add	An S-VLAN tag is attached.
	Service VLAN+customer VLAN (S-VLAN+C-VLAN)	add double	Two VLAN tags, the outer S-VLAN tag and inner C-VLAN tag, are attached.
C-VLAN	C-VLAN	transparent	The VLAN tag is transparently transmitted.
	S-VLAN	translate	One VLAN tag is translated.
	S-VLAN+C-VLAN	add	An S-VLAN tag is attached.
	S-VLAN+C'-VLAN	translate and add	The C-VLAN tag is translated and the S-VLAN tag is attached.

NOTE

- S-Tag: S-VLAN Tag, indicates the service VLAN tag.
- C-Tag: C-VLAN Tag, indicates the user VLAN tag.
- C'-Tag: C'-VLAN Tag, indicates another user VLAN tag.
- untagged: No VLAN tag

Restrictions and Limitations

In home access services, although the CMC supports multiple service VLANs (S-VLANs), all home access services are within the same broadcast domain. Therefore, the home access services do not support VLAN-based broadcast domains.

33.7 EQAM-based Video Technologies

The edge quadrature amplitude modulation (EQAM) feature enables the MA5633 to function as the gateway of a hybrid fiber coaxial (HFC) network connected to an IP network. The MA5633 multiplexes and modulates IP signals into radio frequency (RF) signals so that the data carried in the IP signals can be transmitted over the HFC network. EQAM universally used in this document is also called IPQAM. This section describes the digital video broadcasting (DVB) and video on demand (VoD) services based on the EQAM built in the distributed converged cable access platform (D-CCAP).

NOTE
Only the MA5800 and the MA5633 (cable outlets on a single side of only one cover) support the EQAM-based DVB service.

Why Is Built-in EQAM Required

Traditional EQAM Technology

In multiservice operator (MSO)-oriented solutions, both HFC networks and set top boxes (STBs) use the broadcast television QAM technology and support only RF signals.

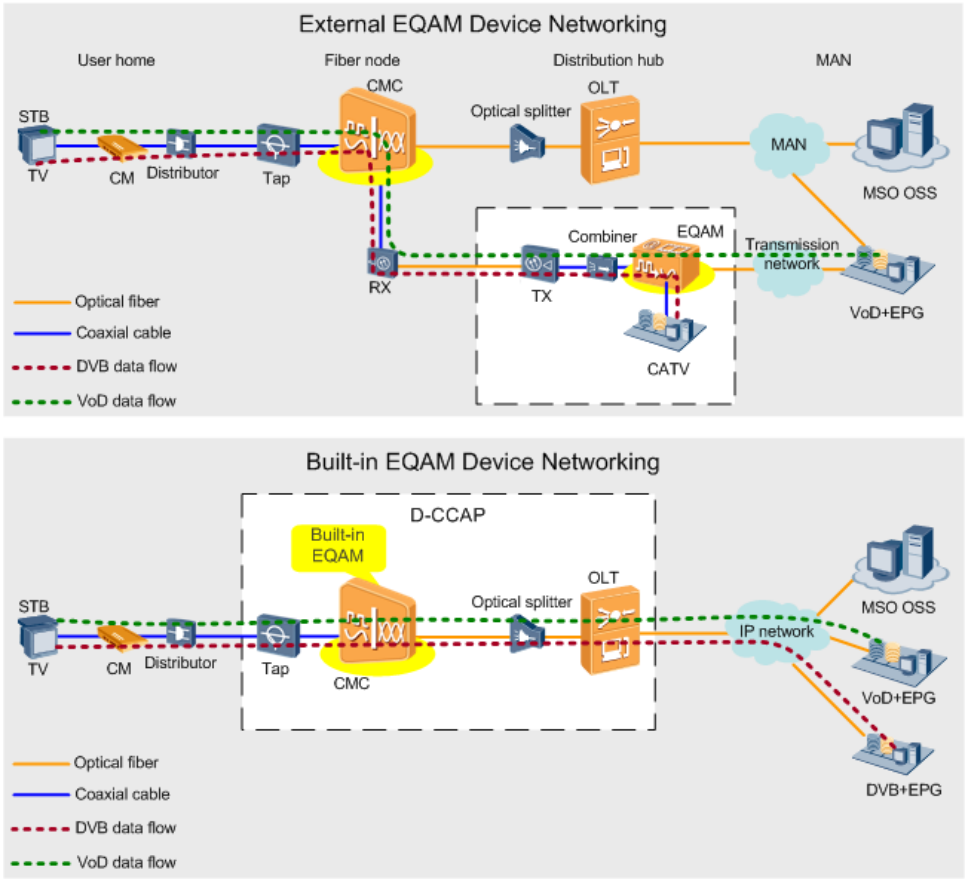
Video data is encapsulated into IP and UDP packet headers that are transmitted on IP networks, which cannot be transmitted on HFC networks. To transmit the video data on HFC networks, the video data must be modulated into RF signals through EQAM.

Video data transmission on HFC networks uses existing cable television (CATV) coaxial cables on the HFC networks, thereby reducing network deployment costs.

EQAM Device Networking Evolution

EQAM devices are classified as external EQAM devices and built-in EQAM devices. Figure 33-28 shows the networking of an external EQAM device and a built-in EQAM device, separately.

Figure 33-28 EQAM device networking



- In **External EQAM Device Networking**, the external EQAM device is deployed at the distribution hub. It modulates video signals into RF signals and sends the RF signals to the STB through the CATV signal channel. The external EQAM device faces the following bottlenecks:
  - In network deployment, the distribution hub is too crowded to support the deployment of an EQAM device.
  - In new site deployment or site reconstruction in both downstream and upstream directions, the environment in the site equipment room is too poor to support the installation of an external EQAM device.
- In **Built-in EQAM Device Networking**, the EQAM device is built in an MA5633. The MA5633 is deployed on a fiber node. It modulates video signals into RF signals and sends the RF signals to the STB through idle downstream channels. The built-in EQAM device has the following advantages:
  - The location of the MA5633 equipped with a built-in EQAM module is moved downwards from the distribution hub to a fiber node, which reduces the space load in the distribution hub.
  - Any idle channel can be used to transmit the video data. Therefore, no dedicated video data transmission channel is required, thereby reducing network construction costs.
  - Different MA5633s can use the same EQAM frequency, improving frequency utilization.

The MA5633 equipped with a built-in EQAM module simplifies the distribution hub and networking complexity as well as reduces maintenance costs. This configuration digitizes MSO networks.

NOTE
The built-in EQAM configuration is used in the remainder of this document, unless otherwise specified.

33.7.2 Basic Concepts

EQAM Management

EQAM Channel

An MA5633 downstream channel can be a data over cable service interface specification (DOCSIS) channel or an EQAM channel.

- A DOCSIS channel carries the data service.
- An EQAM channel carries the video service.

EQAM Channel Profile

An EQAM profile contains common channel parameters for transport streams (TSs), including the transmit period of program association table (PAT) and program map table (PMT) packets as well as TS IDs. This facilitates channel management and maintenance.

Video Profile

In centralized management mode, an optical line terminal (OLT) can manage a maximum of 128 EQAM modules and each EQAM module supports a maximum of 512 video mappings. Then, a large number of (a maximum of 128 x 512) video mappings need to be configured and managed on an OLT.

A video profile can simplify video mapping configuration and management. Specifically, video mappings are configured in the unit of an EQAM module. All video mappings on an EQAM module are defined in a video profile. In addition, the video profile supports batch video mapping configurations.

Multiple EQAM modules can use the same video profile.

Video Domain

A video domain defines the relationships between an EQAM module and an EQAM IP address, and between the service VLAN (S-VLAN) used for video data forwarding on an EQAM module, the EQAM module, and a video profile. This facilitates data management.

Each MA5633 supports only one video domain. In a video domain, an EQAM module maps only one EQAM IP address.

Video

TS

A transport stream (TS) carries one or multiple programs. A single program transport stream (SPTS) carries only one program, and a multiple program transport stream (MPTS) carries multiple programs.

- An SPTS, a TS stream carrying only one program, is used in VoD or switched digital video (SDV), where each program is individually delivered.
- An MPTS, a TS stream carrying multiple programs, is used in broadcast TV (BTV) or SDV, where a pre-bundled set of programs are delivered as a whole.
- VoD is a service which allows users to select their favorite programs from the program list delivered by the digital TV system and watch the programs in real time. When watching programs, users can perform operations, such as pause, fast forward, fast rewind, and locate. Users are generally charged on a monthly or pay-per-view basis.
- The increased number of digital TV programs improves user experience but takes many frequency resources. Specifically, 80% of users watch only 20% of TV programs and the remaining 80% of TV programs are seldom watched. This is long tail. In this case, the bandwidth usage is low for the TV programs that are seldom watched.

  SDV can resolve this issue. Specifically, the SDV system remains the broadcast mode for the TV programs with a rating higher than a specified threshold and uses the VoD mode for other TV programs. The SDV system sends a TV program only to a user group who orders the TV program at a specific time. If the user group switches the TV program, the SDV system replaces the TV program with a new one that has been ordered. In this way, static frequency usage is changed to dynamic frequency usage, thereby improving bandwidth usages.
- BTV is similar to traditional CATV or satellite TV because of the same user experience in watching TV programs, regardless of whether live or traditional TV programs are watched.

Program Elements

Program elements include video streams, audio streams, program specific information (PSI), and data packets, such as program clock reference (PCR) packets.

PSI can be classified into four types:

- PAT: specifies each program number and the PMT packet ID (PID) of a program carried in a TS.
- PMT: specifies video, audio, and PCR PID of a program. All information about a program must be contained in a PMT and a PMT can contain the information about multiple programs.
- Conditional access table (CAT): specifies the encrypted entitlement control message (ECM) and certificate authority (CA) of a specified code stream.
- Network information table (NIT): specifies the channel of a specified TS.

Unscrambled Stream

A video stream that is not encrypted is an unscrambled stream. The VoD head end sends unscrambled stream data to STBs, which features poor security because the STB users can watch programs for free.

Scrambled Stream

The pre-scrambling server encrypts the ordered TV program, generates an encryption file, and stores the file to the video server. When a user orders the TV program, the video server sends the scrambled stream to the user. The STB connected to the user can play the TV program only after obtaining a decryption key, which prevents unpaid users from watching the TV program.

Unscrambled streams differ from scrambled ones only in moving picture experts group (MPEG) control header data. Their configurations are the same on the MA5633 or OLT.

Region ID

A region ID identifies the region of an STB. Video data is encapsulated into UDP packets on IP networks. The destination IP address of the packets is the IP address of the EQAM on the MA5633.

The following section uses an example to describe the function of a region ID:

There are three regions: region 1, region 2, and region 3. Each region is deployed with an MA5633. The three MA5633s use different IP addresses.

If a user in region 1 orders a program, the VoD server must send the program to the MA5633 in region 1 and the destination IP address of the program packet must be the IP address of the MA5633 in region 1. Then, the VoD server must identify the user region and locate the IP address of the MA5633 in this region. Therefore, the STB must obtain its region ID before sending a program ordering request to the MA5633 and VoD server. The STB can obtain its region ID using either of the following solutions:

- The VoD server sends the region ID to the STB. Specifically, the VoD server uses a separate TS to send the region ID to the STB, the MA5633 transparently transmits the TS to the STB, and the STB receives this TS at a specified frequency. This solution is commonly used.
- The MA5633 sends the region ID to the STB. Specifically, the MA5633 sends the TSID configured by running the **stream-id** command to the STB as the region ID.

33.7.3 Principles

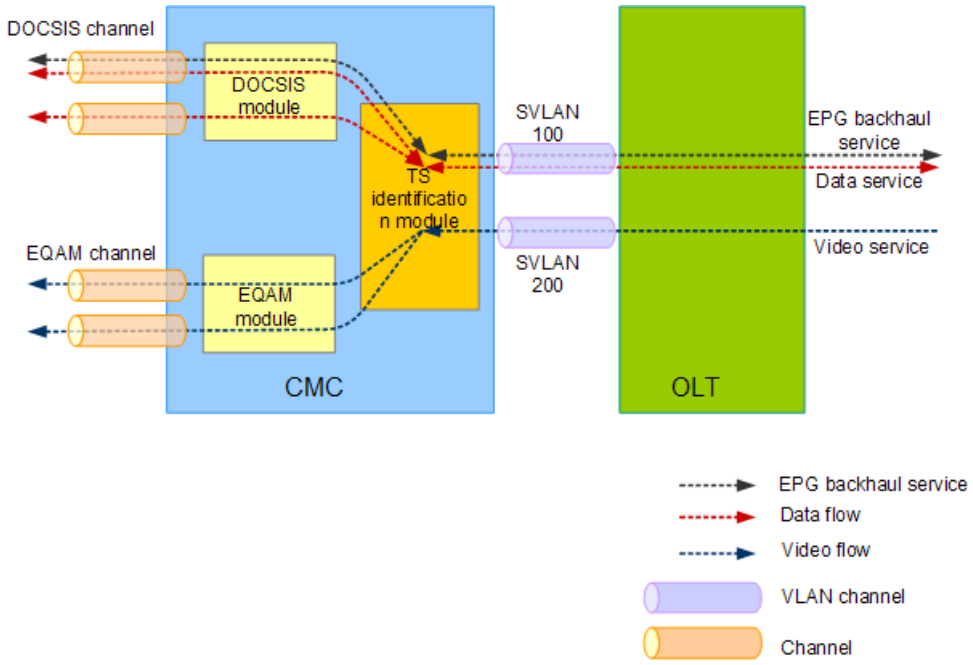
The OLT uses the same channel to send both data and video services to the same uplink port on the MA5633. How does the MA5633 differentiate between the two services?

As shown in Figure 33-29:

- Video packets carry the EQAM IP address and S-VLAN ID. Data packets carry only the S-VLAN ID.
- The TS identification module differentiates between the packets and sends the video packets to the EQAM module for processing. Then, the MA5633 sends the video packets to user-side set top boxes (STBs) through EQAM channels.

  The TS stream identification module sends the data packets to the DOCSIS module for processing (for details, see 33.7.4 Key Technologies for Processing Video Services). Then, the MA5633 sends the data packets to user-side cable modems (CMs) through DOCSIS channels.
- The electronic program guide (EPG) backhaul service of the video service uses the same channel as the data service. It can share service flows with the data service or use different service flows for transmission.

Figure 33-29 Service forwarding principles



33.7.4 Key Technologies for Processing Video Services

Key technologies in video processing include dejittering, multiplexing, straight through, and PCR recovery.

Figure 33-30 Process of implementing the VoD service

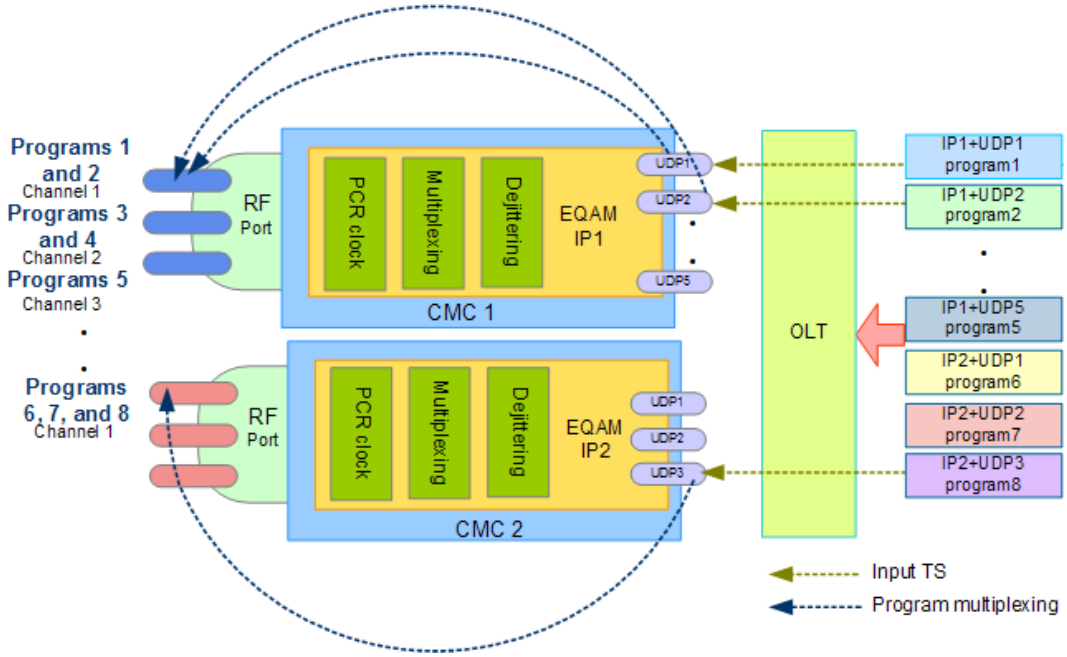
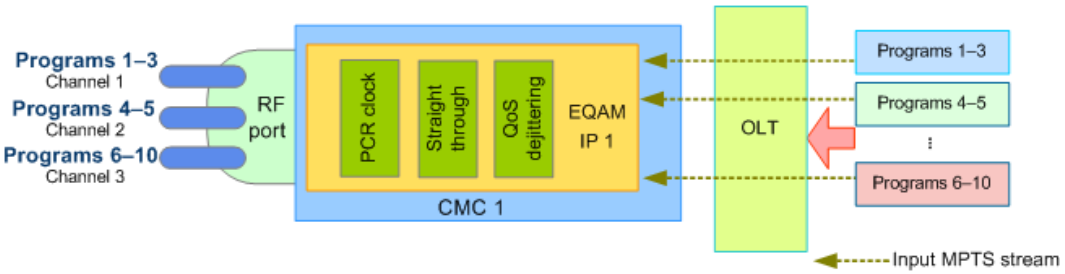


Figure 33-31 Process of implementing the DVB service



Dejittering

The video service sent from a head end is transmitted over IP networks. Delay and jitter are inevitable on IP networks, which degrade the quality of the video service.

Dejittering removes the jitter introduced by the IP networks. This ensures stable conversion from IP signals to RF signals and high quality of IP signal restoration.

The MA5633 supports the configuration of the maximum dejittering size for each SPTS.

Figure 33-32 Dejittering effect

Multiplexing

What is multiplexing?

Multiple separate program streams are combined into an EQAM channel for transmission. This is multiplexing.

Why do we need to multiplex separate program streams?

The following section provides an example to describe the meaning of multiplexing.

An MA5633 RF port is assumed to provide 5 EQAM channels to carry the video service and the users connected to the RF port concurrently order 100 programs. Then, the MA5633 must send the 100 programs to STBs through the 5 EQAM channels. The process that the MA5633 combines the 100 programs into 5 channels is multiplexing.

How is multiplexing implemented?

As shown in Figure 33-30, TSs enter OLT ports. A TS is identified by an different IP address and UDP port number.

An MA5633 provides RF ports on the user side. Program mappings must be configured when the MA5633 is installed and deployed. Program mappings refer to the relationships between the EQAM IP address+UDP port number and an EQAM channel of an RF port. This configuration multiplexes programs.

As shown in Table 33-16, TS 1 is identified by IP 1+UDP 1 and TS 2 is identified by IP 1+UDP 2. Both IP 1+UDP 1 and IP 1+UDP 2 are mapped to channel 1@MA5633 1 in slot 1. After the multiplexing, the TS output from channel 1@MA5633 1 in slot 1 carries programs 1 and 2. The rule applies to other TSs in the table.

Table 33-16 PMT

TSID	IP Address	UDP Port Number	Input Program Number	MA5633 Frame ID in Centralized Management Scenarios	Output Channel ID	Output Program Number
1	10.10.10.10	50	1	1	1	1 and 2
2	10.10.10.10	60	2			
3	10.10.10.10	70	3	1	2	3 and 4
4	10.10.10.10	80	4			
5	10.10.10.10	90	5	1	3	5
6	10.10.20.10	50	6	2	1	6, 7, and 8
7	10.10.20.10	60	7			
8	10.10.20.10	70	8			

Straight Through

As shown in Figure 33-31, input MPTS streams are simply output from specified channels, and the programs carried over each MPTS stream remain unchanged.

PCR Recovery

In the video service, video and audio data are transmitted using different packets. When a video is played, the video and audio data must match. The PCR recovery technology ensures synchronous video and audio playing.

33.7.5 Networking Applications

The typical application of the built-in EQAM module is to support the VoD service.

As shown in Figure 33-33:

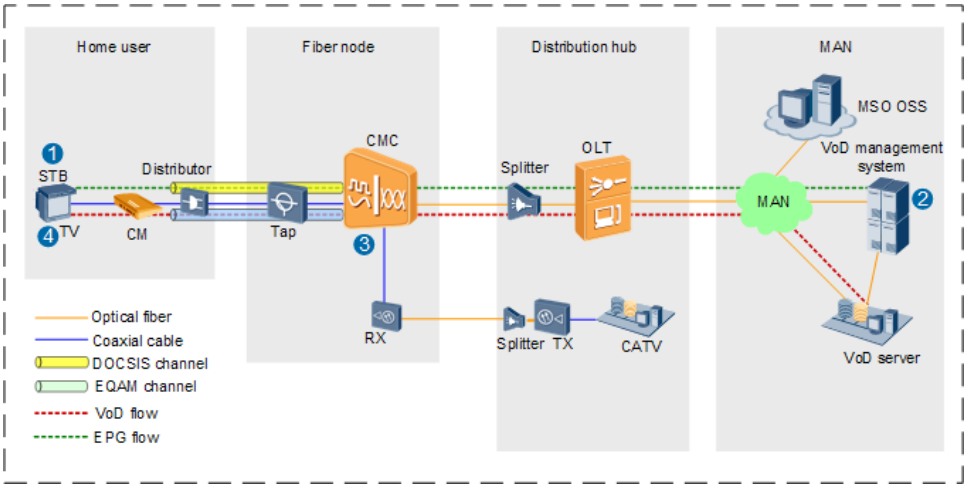
- The VoD management system manages users, stores the EPG, and implements charging.
- The VoD server stores programs.

The process of ordering a program is as follows:

1. An STB user orders a program from the EPG.

   The packets involved in this process are carried over DOCSIS channels.
2. The VoD management system authorizes the user and instructs the VoD server to send the VoD data to the MA5633 through the IP network.
3. The EQAM module built-in the MA5633 multiplexes and modulates the IP data into RF signals and then sends the RF signals to the STB through EQAM channels over the HFC network.
4. The STB plays the program after decrypting the RF signals.

Figure 33-33 VoD networking



33.7.6 Configuring EQAM

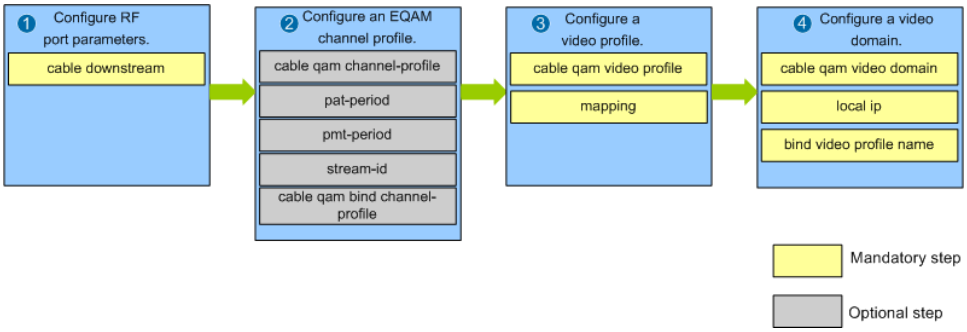
Prerequisites

The ARP sending period of the OLT's upper-layer gateway must be shorter than the MAC address aging time of the forwarding device (OLT or CMC) on the link. Otherwise, the packets forwarded by the upper-layer gateway to the OLT or CMC will be changed to unknown unicast packets, and packet loss will occur, leading to artifacts on the screen during the display of a video.

Configuration Process

Figure 33-34 shows the process of configuring EQAM.

Figure 33-34 Process of configuring EQAM



Procedure

Run the **cable downstream** command to configure RF port parameters.

Configuring RF port parameters implements video data transmission over cables. RF port parameters include:

- **frequency**: indicates the center frequency. A center frequency and a frequency bandwidth determine a frequency range. Packets are transmitted within the frequency range.
- **modulation**: indicates the modulation mode of downstream channels. A greater **modulation** value results in a higher transmission bandwidth but a less stronger anti-interference capability. Therefore, configure the **modulation** value based on line conditions.
- **channel-mode**: indicates a channel mode. The default mode is DOCSIS. A channel can work only in one channel mode.
- **symbol-rate**: indicates the symbol rate, data volume transmitted within a specific period of time. The **symbol-rate** value is determined based on signal bit rates and channel parameter settings.

Step 1 Configure an EQAM channel profile.

An EQAM profile contains channel parameters for transport streams (TSs).

1. Run the **cable eqam channel-profile** command to create an EQAM channel profile.
2. Run the **pat-period** command to configure the period of transmitting PAT packets.

PAT: specifies each program number and the PMT packet ID (PID) of a program carried in a TS.

3. Run the **pmt-period** command to configure the period of transmitting PMT packets.

   PMT: specifies video, audio, and program clock reference (PCR) PID of a program.

4. Run the **stream-id** command to configure a TSID.
5. Run the **cable eqam bind channel-profile** command to bind the EQAM channel profile to a specified RF port.

Step 2 Configure a video profile.

A video domain defines the relationships between an EQAM module and an EQAM IP address, and between the service VLAN (S-VLAN) used for video data forwarding on an EQAM module, the EQAM module, and a video profile. This facilitates data management.

1. Run the **cable eqam video profile** command to create a video profile.
2. Run the **mapping** command to configure program mapping.

   The configuration of program mapping implements multiplexing, which combines and sends multiple programs over one channel to an STB connected to a CM.

   Program mapping involves the following parameters:

   - **udp-port**: indicates the number of a UDP port through which the video data is input from the network side. No well-known port numbers are allowed.
   - **jitter**: indicates a network jitter that is removable on the MA5633.
   - **pmtpid**: indicates the PMT PID of an output program. A PMT PID must be unique in an EQAM channel and can be repetitive in different EQAM channels.
   - **output-program**: indicates an output program number. An output program number can retain the input program number or be reconfigured. Ensure that an output program number is unique in an EQAM channel.

Step 3 Configure a video domain.

1. Run the **cable eqam video domain** command to create a video domain.
2. Run the **local ip** command to configure the IP address and VLAN ID of the video domain.
3. Run the **bind video profile name** command to bind the video profile to a specified RF port in the video domain.

----End

Example

The following is an example of the configurations used to enable EQAM for a new MA5633 network that covers a few number of users:

NOTE

In this example, the MA5633 works in centralized management mode. In standalone NE mode, you must only change the port number to 0/1/0.

- Number of DOCSIS channels: 12
- Number of EQAM channels: 4
- EQAM channel ID: 9-12
- Modulation mode: QAM 256 and QAM 64
- Symbol rate: 6.875 or 6.900

- UDP port number: 501-514

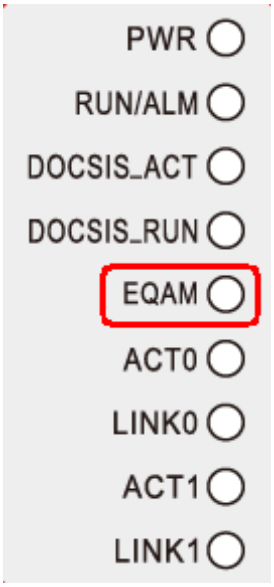
```
huawei(config)#interface cable 1/1/0
huawei(config-if-cable-1/1/0)#cable downstream 9 frequency 115.0 enable channel-mode eqam symbol-rate 6.875
huawei(config-if-cable-1/1/0)#cable downstream 10 frequency 123.00 enable channel-mode eqam symbol-rate 6.900
huawei(config-if-cable-1/1/0)#cable downstream 11 frequency 131.00 enable channel-mode eqam symbol-rate 6.900
huawei(config-if-cable-1/1/0)#cable downstream 12 frequency 139.00 modulation qam64 enable channel-mode eqam symbol-rate 6.875
huawei(config-if-cable-1/1/0)#quit
huawei(config)#cable eqam video profile name vod1
huawei(config-eqam-video-profile-vod1)#mapping index 1 type spts udp-port 501 output-program 1 pmtpid 32 channel 9
huawei(config-eqam-video-profile-vod1)#mapping index 2 type spts udp-port 502 output-program 2 pmtpid 48 channel 9
huawei(config-eqam-video-profile-vod1)#mapping index 3 type spts udp-port 503 output-program 3 pmtpid 64 channel 9
huawei(config-eqam-video-profile-vod1)#mapping index 4 type spts udp-port 504 output-program 4 pmtpid 80 channel 9
huawei(config-eqam-video-profile-vod1)#mapping index 5 type spts udp-port 505 output-program 1 pmtpid 96 channel 10
huawei(config-eqam-video-profile-vod1)#mapping index 6 type spts udp-port 506 output-program 2 pmtpid 112 channel 10
huawei(config-eqam-video-profile-vod1)#mapping index 7 type spts udp-port 507 output-program 3 pmtpid 128 channel 10
huawei(config-eqam-video-profile-vod1)#mapping index 8 type spts udp-port 508 output-program 1 pmtpid 160 channel 11
huawei(config-eqam-video-profile-vod1)#mapping index 9 type spts udp-port 509 output-program 2 pmtpid 176 channel 11
huawei(config-eqam-video-profile-vod1)#mapping index 10 type spts udp-port 510 output-program 1 pmtpid 224 channel 12
huawei(config-eqam-video-profile-vod1)#mapping index 11 type spts udp-port 511 output-program 2 pmtpid 240 channel 12
huawei(config-eqam-video-profile-vod1)#mapping index 12 type spts udp-port 512 output-program 3 pmtpid 256 channel 12
huawei(config-eqam-video-profile-vod1)#mapping index 13 type spts udp-port 513 output-program 4 pmtpid 272 channel 12
huawei(config-eqam-video-profile-vod1)#mapping index 14 type spts udp-port 514 output-program 5 pmtpid 288 channel 12
huawei(config-eqam-video-profile-vod1)#quit
huawei(config)#cable eqam video domain huawei
huawei(config-eqam-domain-huawei)#local ip 10.10.10.10 vlan 45
//The VLAN has been created.
huawei(config-eqam-domain-huawei)#bind video profile name vod1 1/1/0
huawei(config-eqam-domain-huawei)#quit
```

33.7.7 Maintenance and Diagnosis

EQAM Indicator

The MA5633 provides an EQAM indicator to show the EQAM status.

Figure 33-35 Indicator



NOTE
The MA5633 with the dimensions of 189 mm x 308 mm x 153 mm (H x W x D) provides the EQAM indicator.

Status	Description
Off	EQAM is disabled.
Orange, on for 0.125s and off for 0.125s repeatedly	EQAM is faulty.
Steady green	EQAM is functional and no data is transmitted.
Green, on for 0.5s and off for 0.5s repeatedly	EQAM is functional and data is being transmitted.

EQAM Packet Statistics

The MA5633 supports packet statistics query. Based on the statistical results, you can determine whether users can order programs and whether packet loss occurs in program transmitting.

- Run the **display cable eqam video stream** command to check the TS status. If the status is **online**, the TS is functional.
- Run the **display cable eqam video statistics input** command to query packet statistics and network jitter of a TS on the video service input end.

  Check whether the program ordered by a user has been sent to the MA5633. If the program has been sent to the MA5633, check whether the program traffic is normal.
- Run the **display cable eqam video statistics output** command to query packet statistics of a TS on the video service output end.

Check whether the TS contains the ordered program, whether the program has been sent to the STB, and whether the TS traffic is normal.

Event

An event reflects EQAM channel status. If an event is reported, handle the event to eliminate faults.

Event	Name
0x66300003	The EQAM channel utilization exceeds the threshold

EQAM Ping

When the video service fails, ping the EQAM IP address from an OLT or an upper-layer device to check the network connection. If the ping fails, the MA5633 is disconnected from the upper-layer device. In this case, check network route settings.

- An EQAM module can only be pinged.
- The EQAM IP address cannot be pinged from the user side.
- In centralized management mode, if the OLT pings the MA5633, the MA5633 replies with an ARP response only after the RF port is enabled.
- The MA5633 supports only ICMP Echo Reply packets.
- The MA5633 can process the ICMP Request packets only with a length less than 1500 bytes.

33.7.8 Standards and Protocols Compliance

Standard or Protocol	Description
ISO/IEC 13818-1	Information Technology-Generic Coding of Moving Picture and Associated Audio: system
ISO/IEC 13818-2	Information Technology-Generic Coding of Moving Picture and Associated Audio: video
ISO/IEC 13818-3	Information Technology-Generic Coding of Moving Picture and Associated Audio: audio
J.83A	Digital multiprogrammed systems for television, audio, and data services for cable distribution

33.8 Load Balancing

The rapid development of broadband services and continuous user increasing promote high requirements on network bearing capabilities for cable carriers. The cable carriers consistently concern about broadband quality of service (QoS) and port traffic. Network infrastructure

reconstruction and port capacity expansion can resolve traffic saturation at the access layer. However, these operations require not only heavy engineering workload but also high costs.

The load balancing feature enables the distributed converged cable access platform (D-CCAP) to improve broadband QoS without requiring network infrastructure reconstruction, thereby significantly reducing engineering workload and costs.

What Is Load Balancing

A D-CCAP provides multiple downstream and upstream channels for radio frequency (RF) ports. The cable modems (CMs) connected to the D-CCAP share all the channels. A CM can randomly use downstream and upstream channels to go online, which may cause load unbalancing on channels and consequently packet loss. The load balancing feature allows the D-CCAP to migrate the CMs on a heavy-load channel to a light-load channel, which balances loads between channels and maximally uses RF channel bandwidths.

Load Balancing Group

The D-CCAP supports two types of load balancing groups: a general load balancing group and multiple restricted load balancing groups. The D-CCAP adds CMs to a general or restricted load balancing group to implement different load balancing policies.

- General load balancing group

  A general load balancing group is created by the D-CCAP by default. It is open to all the downstream and upstream channels with load balancing enabled. Each D-CCAP supports only one general load balancing group.

  A general load balancing group shares channel resources and does not differentiate between CMs or service types. It is used for common user services.

- Restricted load balancing group

  The CMs of VIP users added to a restricted load balancing group preferentially use channel resources in this group. A restricted load balancing group is manually added. In a restricted load balancing group, the downstream and upstream channels as well as the CMs that can be added to the group are specified.

  Users can add CMs to one restricted load balancing group based on:

  - MAC addresses in a specific range
  - Data over cable service interface specification (DOCSIS) versions. For example, add all DOCSIS 2.0-compliant CMs to one restricted load balancing group.
  - Service types or other user characteristics

A CM can be added to only one load balancing group. All CMs are added to the general load balancing group by default. If a CM meets the requirements of a restricted load balancing group, it is preferentially added to this group.

33.8.2 Load Balancing Types

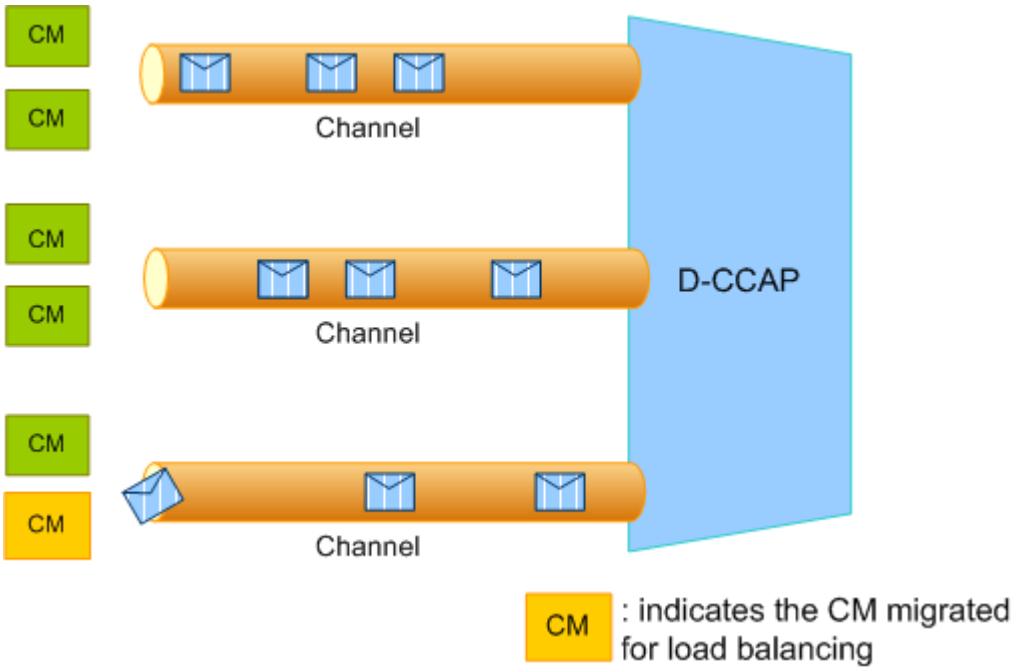
The D-CCAP balances load based on the number of CMs, service flow bandwidths, or bandwidth usages.

Based on the Number of CMs

The D-CCAP evenly distributes CMs to multiple channels based on the number of CMs on each channel.

Application Scenario: Service bandwidths of CMs are balanced. This load balancing mode is simple to implement and applies to simple service scenarios.

Figure 33-36 Load balancing based on the number of CMs

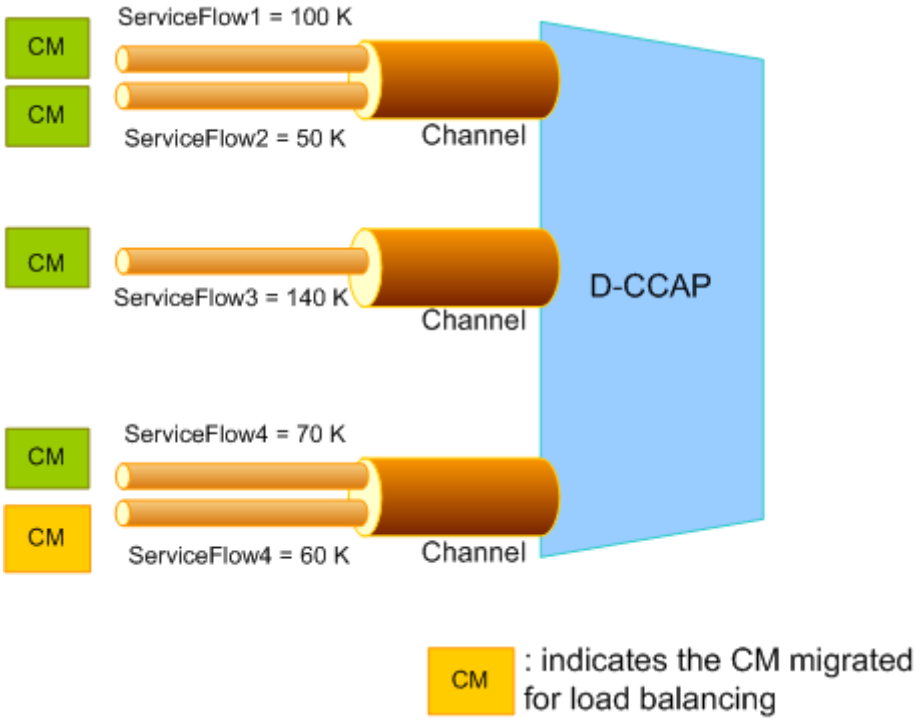


Based on Service Flow Bandwidths

The D-CCAP evenly distributes CMs to multiple channels based on service flow bandwidths. The sum of the minimum assured bandwidths for online CMs on a channel is assumed to be value A, and the theoretical bandwidth of the channel is assumed to be value B. The D-CCAP compares the ratios of value A to value B between channels and ensures that the load of each channel is balanced. When a new CM goes online, the D-CCAP distributes the CM to the channel with the lowest ratio of value A to value B.

Application Scenario: Service bandwidths of each CM are unbalanced, and the D-CCAP needs to precisely distribute the CMs based on service flow bandwidths. This load balancing mode applies to complex service scenarios.

Figure 33-37 Load balancing based on service flow bandwidths



This load balancing type involves the following basic concepts:

- Static service flow bandwidth statistics for channels = Minimum assured bandwidth of the main service flow/Number of channels

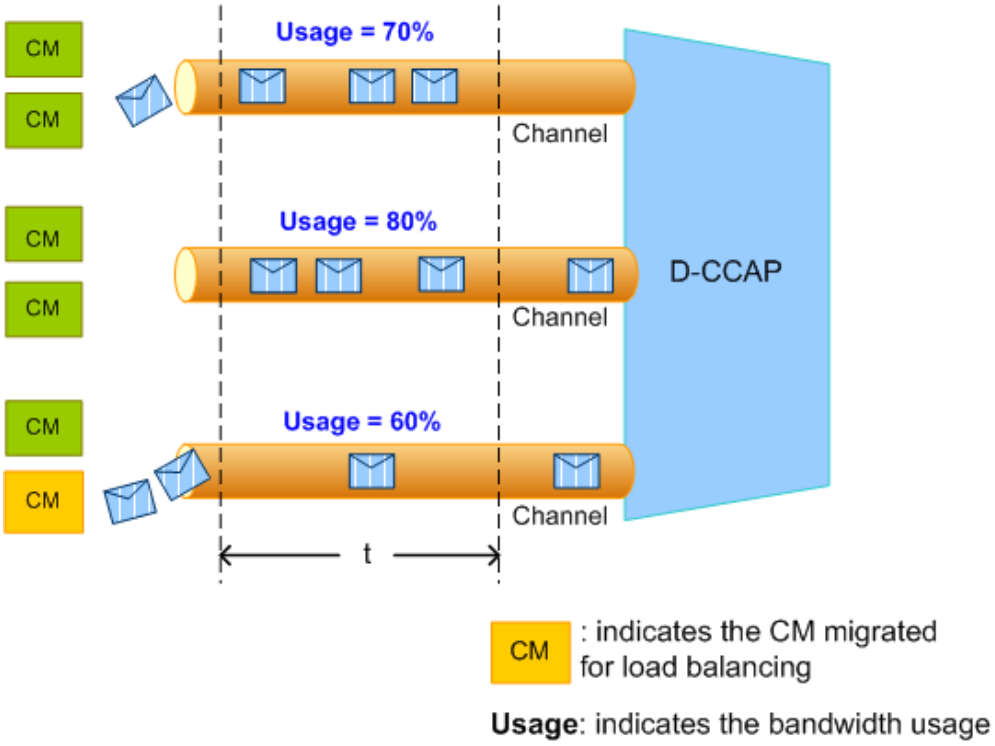
  The D-CCAP collects the minimum assured bandwidth of only the main service flow. If no minimum assured bandwidth has been configured for the main service flow, the D-CCAP uses default value 100 kbit/s for the main service flow.
- Service flow bandwidth statistics for each channel = Minimum assured bandwidth of the main service flow/Number of channels

Based on Bandwidth Usages

After the period of collecting channel bandwidth usages ends, the D-CCAP balances the load of the channels used by CMs if the deviation of the bandwidth usages between channels exceeds the threshold. The D-CCAP uses this load balancing type only when the total bandwidth usage is greater than 20%. If the total bandwidth usage is not greater than 20%, the D-CCAP balances load based on service flow bandwidths, even if the load balancing type configured on the D-CCAP is based on bandwidth usages.

Application Scenario: Service bandwidths of each CM are unbalanced, and the D-CCAP needs to balance network load in real time. This load balancing mode applies to scenarios with high requirements for load balancing.

Figure 33-38 Load balancing based on bandwidth usages



This load balancing type involves the following basic concepts:

- Dynamic bandwidth usage of a channel

  Due to traffic burst, the dynamic bandwidth usage of a channel may be high in a period of time, which cannot reflect stable channel load. Therefore, the channel load must be considered in multiple bandwidth usage statistical periods. The D-CCAP calculates the bandwidth usage of a channel using the calculation methods listed in the following table.

Table 33-17 Calculation method

Number of Bandwidth Usage Statistical Periods	Formula
1	Calculated dynamic bandwidth usage = Bandwidth usage in the current period
2	Calculated dynamic bandwidth usage = (3 x Bandwidth usage in the current period + Bandwidth usage in the previous period)/4
4	Calculated dynamic bandwidth usage = (9 x Bandwidth usage in the current period + 4 x Bandwidth usage in the previous period + 2 x Bandwidth usage in the previous two periods + Bandwidth usage in the previous three periods)/16

- Dynamic bandwidth usage of an upstream channel for a CM

  The D-CCAP obtains the dynamic bandwidth used by a CM on an upstream channel based on timeslot allocation on the upstream channel. The D-CCAP obtains the values in multiple bandwidth usage statistical periods and calculates the dynamic bandwidth usage of the upstream channel for the CM. The method of calculating the dynamic bandwidth usage of an upstream channel for a CM is the same as that of calculating the dynamic bandwidth usage of a channel. For details, see Table 33-17.

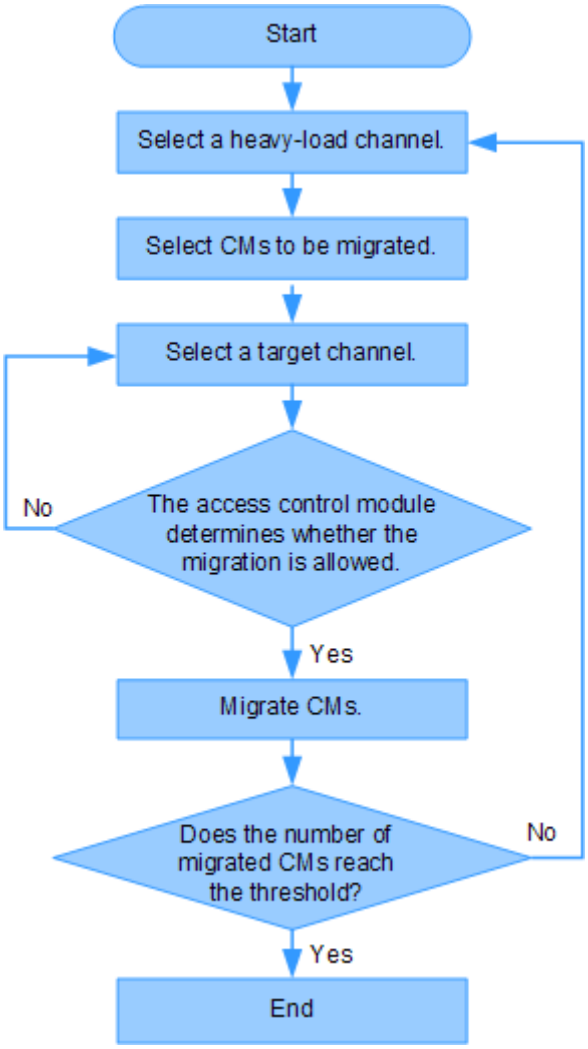
- Dynamic bandwidth usage of a downstream channel for a CM
  - For a DOCSIS 2.0-compliant CM, the D-CCAP calculates the dynamic bandwidth usage of a downstream channel for the CM based on bandwidth statistics for each service flow of the CM.
  - A DOCSIS 3.0-compliant CM can use multiple downstream channels. The bandwidth for each service flow of a DOCSIS 3.0-compliant CM is allocated in multiple downstream channels based on the bandwidth usages of the downstream channels. The formula for calculating the dynamic bandwidth usage of a downstream channel for a DOCSIS 3.0-compliant CM is as follows:

    Dynamic bandwidth usage of a downstream channel for the CM = The bandwidth used by each service flow of this CM x (Used channel bandwidth - Bandwidth used by DOCSIS 2.0-compliant CMs)/(Total used bandwidth of all the channels used by DOCSIS 3.0-compliant CMs - Total bandwidth used by DOCSIS 2.0-compliant CMs on these channels)
  - The D-CCAP obtains the values in multiple bandwidth usage statistical periods and calculates the dynamic bandwidth usage of a downstream channel for the CM. The method of calculating the dynamic bandwidth usage of a downstream channel for a CM is the same as that of calculating the dynamic bandwidth usage of a channel. For details, see Table 33-17.

33.8.3 Load Balancing Process

During the load balancing process, the D-CCAP periodically migrates some CMs from heavy-load channels to light-load channels based on the loads on each channel for balancing loads between channels.

Figure 33-39 Load balancing flowchart



Periodically Starting Load Balancing

The D-CCAP periodically starts load balancing based on channel loads.

Selecting a Heavy-Load Channel

- Load balancing based on the number of CMs:
  a. The D-CCAP calculates the average number of CMs on channels.
  b. The D-CCAP obtains a channel list for balancing loads by sorting the channels with the number of CMs exceeding the average number in descending order.
- Load balancing based on service flow bandwidths and bandwidth usages:
  a. The D-CCAP calculates the average bandwidth usage of channels. Then, the D-CCAP uses the larger value between the average value and the value of parameter **trigger** (specifying the threshold for triggering load balancing) as the threshold for balancing loads for channels.
  b. The D-CCAP obtains a channel list for balancing loads by sorting the channels with bandwidth usages exceeding the load balancing threshold in descending order.

NOTE

The differences between the two load balancing types are as follows:

- For load balancing based on service flow bandwidths, the D-CCAP uses the static bandwidth usage of the minimum assured bandwidth for the service flow to calculate the bandwidth usage of each channel.
- For load balancing based on bandwidth usages, the D-CCAP uses the actual dynamic bandwidth usage to calculate the bandwidth usage of each channel.

Selecting CMs to Be Migrated

The D-CCAP selects the CMs to be migrated from the channels in a channel list. The CMs to be migrated must meet the following requirements:

- The CMs support load balancing.
- There is no ongoing voice service on the CMs.
- The duration from the time when the CM is migrated last time to the current time is longer than the value of parameter **cm-move-interval** (specifying the minimum duration).
- For the load balancing based on service flow bandwidths and bandwidth usages, the dynamic bandwidth usage for the CM must be greater than 1% of the dynamic channel bandwidth usage.

Selecting a Target Channel

The general rules of selecting a target channel are as follows:

- If the CM to be migrated is added to a restricted load balancing group, the target channel of this CM must also be contained in this restricted load balancing group.
- If a CM is not added to a restricted load balancing group, all channels can be used as the target channel of this CM.
- The target channel must meet the following requirements:
  a. The loads on the channel are the lightest.
  b. The load difference between the target channel and the source channel must be greater than the value of parameter **diff** (threshold for channel bandwidth usage differences) or the value of parameter **num-diff** (threshold for CM quantity differences).
- For DOCSIS 3.0-compliant CMs, the loads on the target channel set must be lightest and the channel IDs must be continuous.

Controlling Admission

The D-CCAP implements admission control based on service flow types during load balancing. For admission control concepts, see 33.9 Admission Control.

Before the D-CCAP migrates a CM to the target channel, the access control module checks the request of a CM service flow to determine whether the CM can use the target channel to go online.

- If the request meets requirements, the D-CCAP migrates the CM to the target channel.
- If the request does not meet requirements, the D-CCAP needs to select an appropriate target channel for the CM again. If no target channel is available, the D-CCAP does not migrate the CM.

Migrating CMs

- For DOCSIS 2.0-compliant CMs, the D-CCAP sends dynamic channel change (DCC) messages to the CMs for migration.
- For DOCSIS 3.0-compliant CMs, the D-CCAP sends dynamic bonding-channel change (DBC) messages to the CMs for migration.

Configuring the Maximum Number of CMs That Can Be Migrated in a Period

If the number of CMs to be migrated in a period is large, the CMs may fail to migrate and they go offline. To prevent such an issue from occurring, the maximum number of CMs that can be migrated (**max-move**) in a period must be configured. If the number of migrated CMs in a period reaches this configured value, the D-CCAP stops balancing loads.

Balancing Loads on Main Channels

During CM ranging, the D-CCAP selects the channel with light loads (a small bandwidth usage) as the main channel. This ensures load balancing between the main channel and other channels after the CM goes online.

During periodic load balancing, if the main channel of a CM is contained in the target channels for load balancing, the main channel of the CM remains unchanged. Otherwise, the channel with the minimal number of CMs is used as the main channel of this CM.

33.8.4 Configuring Load Balancing

Load balancing implements balanced traffic transmission over channels. A D-CCAP automatically allocates the traffic over each physical channel based on network conditions, implementing load balancing in a channel group.

Service Requirements

A D-CCAP connects to 10 CMs, CMs 1 through 10.

- CMs 1 and 2 need to be added to a restricted load balancing group to ensure their QoS.
- Load balancing is not implemented on CM 10.
- Other CMs are added to the general load balancing group.

Data Plan

Table 33-18 Data plan for key parameters

Item	Data
Load balancing type	Based on service flow bandwidths
Restricted load balancing group where CMs 1 and 2 are added	• ID of the group: 1 • Channels added to this group: downstream channels 2 and 3 and upstream channel 1
Group where CM 10 is added to and load balancing is not implemented	ID of the group: 10

Procedure

Globally enable load balancing.

```
huawei(config)#cable load-balance enable
```

Step 1 Enable the general load balancing group and configure load balancing based on service flow bandwidths.

The general load balancing group is enabled by default.

```
huawei(config)#interface cable 1/1/0
huawei(config-if-cable-1/1/0)#cable load-balance-group general method service-flow
```

Step 2 Add restricted load balancing group 1, add a downstream channel list and upstream channel list to this group, and add CMs 1 and 2 to this group.

Ensure that CMs 1 and 2 can transmit data over the channels in the restricted load balancing group. The MAC address range of CMs 1 and 2 is as follows: The MAC address is 0010-01ab-0000 and the mask of the MAC address is ffff-ffff-0000.

```
huawei(config-if-cable-1/1/0)#cable load-balance-group restrict add 1
huawei(config-if-cable-1/1/0)#cable load-balance-group channel add 1 downstream 2,3
upstream 1
huawei(config-if-cable-1/1/0)#cable load-balance-group modem add 1 mac 0010-01ab-0000
mask ffff-ffff-0000 group 1
huawei(config-if-cable-1/1/0)#cable load-balance-group restrict enable 1
```

Step 3 Configure CM 10 where load balancing is not implemented.

The MAC address range of CM 10 is as follows: The MAC address is 00e0-fcab-0000 and the mask of the MAC address is ffff-ffff-0000.

NOTE

Load balancing changes the downstream or upstream channel of CMs and accordingly interrupts CM services. Determine the CMs where load balancing is not implemented if load balancing is not required on the CMs or the CMs that do not comply with DOCSIS standards because the CMs may go offline after their channels are changed.

```
huawei(config-if-cable-1/1/0)#cable load-balance-group exclude-modem add 1 mac
00e0-fcab-0000 mask ffff-ffff-0000
```

----End

33.9 Admission Control

Admission control is a mechanism for managing admission requests from service flows when MA5600T/MA5603T/MA5608T resources cannot meet the requirements for registering cable modems (CMs) or dynamically creating service flows.

What Is Admission Control

Introduction

The admission control feature enables the distributed converged cable access platform (D-CCAP) to:

- Prevent service exceptions due to resource exhaustion while providing quality of service (QoS) guarantees.
- Reserve bandwidth for emergency calls to ensure that these calls are given the highest priority.

Admission control is a method of improve the effectiveness of QoS on created service flows, but it is not a method of applying QoS to service flows. That is, admission control controls service flow creation, while QoS is a quality assurance method for created service flows.

The D-CCAP supports admission control based on bandwidth usages of upstream and downstream channels. Admission control can be implemented based on flow characteristics, such as the service class name of service flows, scheduling type of upstream service flows, and priorities of PacketCable service flows. This allows the D-CCAP to implement fine-grained bandwidth management and ensures bandwidths for various service flows. The D-CCAP controls the following requests:

- CM registration
- Dynamic service flow creation, such as for PacketCable calls

Benefits

Admission control enabled on a D-CCAP ensures the normal running of services on this D-CCAP for good user experience.Figure 33-40 and Figure 33-41 show user experience before and after admission control is enabled on a D-CCAP, respectively.

Figure 33-40 Before admission control is enabled

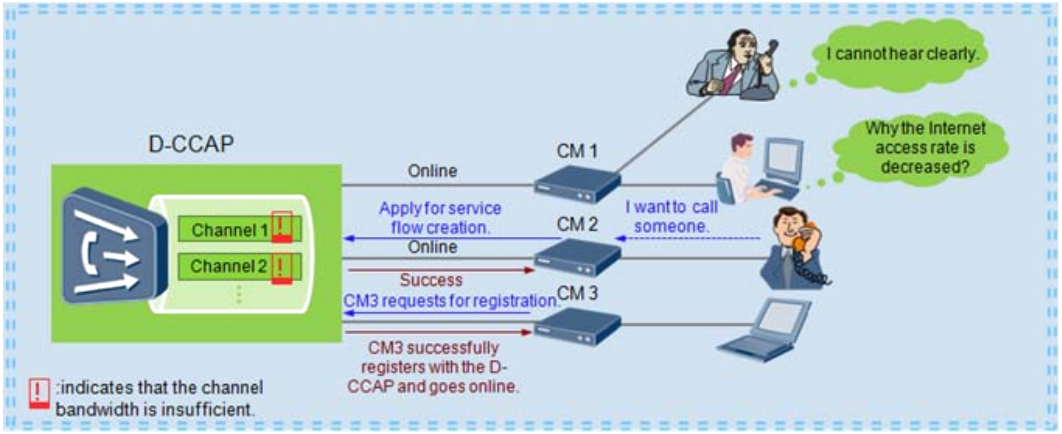
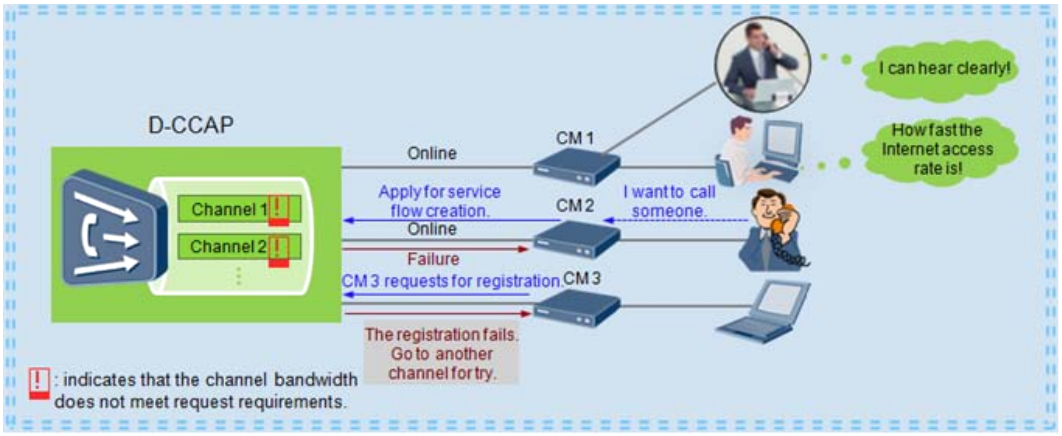


Figure 33-41 After admission control is enabled



33.9.2 Basic Admission Control Concepts

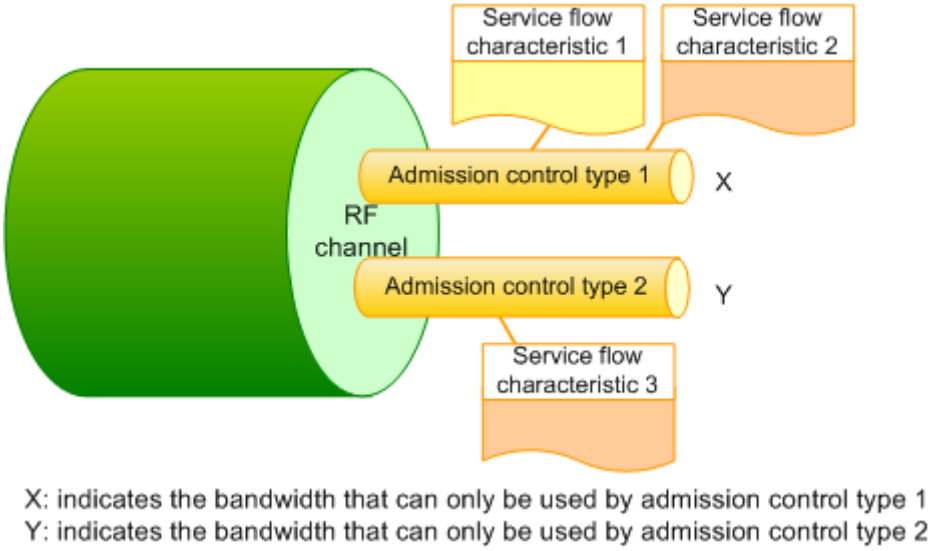
This section describes basic admission control concepts, including service flow characteristics, admission control types, minimum assured bandwidth, and bandwidth threshold.

Overview

Admission control allows the D-CCAP to map service flow characteristics, such as the service class name of service flows, priorities and application types of PacketCable service flows, and scheduling type of upstream service flows, to an admission control type and configure a bandwidth threshold for an admission control type for fine-grained service flow management. As shown in Figure 33-42, two admission control types 1 and 2 are configured for the RF channel.

- Service flow characteristics 1 and 2 are mapped to admission control type 1. The exclusive bandwidth allocated by the D-CCAP to admission control type 1 is X.
- Service flow characteristic 3 is mapped to admission control type 3. The exclusive bandwidth allocated by the D-CCAP to admission control type 2 is Y.

Figure 33-42 Relationships between admission control types and service flow characteristics



Service Flow Characteristics

Service flows can be classified according to the following service flow characteristics:

- Service class name used by a CM for registration
- **Session Class** field specified in the Common Open Policy Service (COPS) protocol

  This field is used to differentiate between common VoIP sessions and high-priority VoIP sessions for PacketCable 1.x service flows.
- **Application Type** or **SessionClassID** field specified in the COPS protocol

  The **Application Type** field is used to differentiate application types and the **SessionClassID** is used to differentiate session priorities for PacketCable Multimedia (PCMM) service flows.
- Service flow scheduling types:
  - Unsolicited grant service (UGS)
  - Unsolicited grant service with activity detection (UGS-AD)
  - Real-time polling service (rtPS)
  - Non-real-time polling service (nrtPS)
  - Best effort (BE)

Admission Control Types

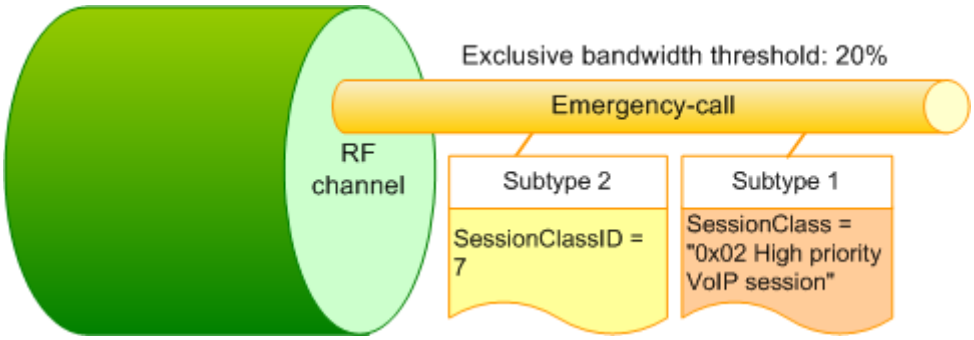
All service flow characteristics are mapped to an admission control type, and a bandwidth threshold is configured for each admission control type for flexible bandwidth management. An admission control type can contain multiple subtypes, which correspond to different flow characteristics. A subtype must be unique among all admission control types.

The following example explains the relationships between bandwidth and service flow characteristics. This example assumes that a 10% exclusive bandwidth of the total bandwidth is reserved for emergency calls and the service flows meeting emergency call requirements have two characteristics: characteristic 1 and characteristic 2.

- Characteristic 1: The **Session Class** field in COPS identifies high-priority VoIP sessions for PacketCable 1.x service flows.
- Characteristic 2: The priority is 7 specified by the **SessionClassID** field in COPS for PacketCable Multimedia service flows.

If a 10% bandwidth threshold is configured for each characteristic, the reserved bandwidths cannot be shared by emergency calls with both characteristics. If the two characteristics are mapped to an admission control type and a 20% exclusive bandwidth is configured for the admission control type, all emergency calls from the service flows with the two characteristics can share the 20% exclusive bandwidth. As shown in Figure 33-43, the admission control type of emergency calls is **Emergency-call** and the exclusive bandwidth threshold is 20%.

Figure 33-43 Admission control for emergency calls



Minimum Assured Bandwidth

A minimum assured bandwidth is a minimum bandwidth required for creating a service flow, which is calculated as follows:

- The formula for calculating the minimum assured bandwidth for UGS and UGS-AD upstream service flows is as follows: Minimum assured bandwidth = Grant size x Number of grants per second.
- The minimum assured bandwidth for other types of upstream and downstream service flows is configured through QoS parameters.

Bandwidth Threshold

The D-CCAP implements admission control by comparing the RF channel bandwidths used by all service flows with the configured bandwidth threshold. A bandwidth threshold consists of an exclusive bandwidth threshold and a non-exclusive bandwidth threshold.

- Exclusive bandwidth is the bandwidth that can only be used by the service flows of a specified admission control type.
- Non-exclusive bandwidth is the remaining bandwidth that is not configured as exclusive bandwidth. The non-exclusive bandwidth can be used by all types of service flows. The non-exclusive bandwidth is used on a first come, first serve basis. If the non-exclusive bandwidth is used by a type a service, other services can use the non-exclusive bandwidth only after the original service releases this bandwidth.
- The sum of exclusive bandwidths and non-exclusive bandwidths configured for the service flows of all admission control types in radio frequency (RF) downstream or upstream channels cannot be greater than the total bandwidth of these channels.

- After the exclusive bandwidth threshold and non-exclusive bandwidth threshold are configured for an admission type, the maximum bandwidth threshold for the service flows of this admission control type is the sum of the exclusive bandwidth threshold and the non-exclusive bandwidth threshold. If unused non-exclusive bandwidths of this admission control type are available, these bandwidths can be used by other services. The configuration of a non-exclusive bandwidth threshold for an admission control type effectively prevents a service from preempting non-exclusive bandwidths, which causes the failure of other services.

33.9.3 How Is Admission Control Implemented

The MA5633 implements admission control for service flows in both centralized management and standalone NE modes. In the remainder of this document, a D-CCAP is an MA5633.

Admission Control Process

As shown in Figure 33-44, three admission control types (types 1, 2, and 3), are configured for the RF channel. The exclusive and non-exclusive bandwidths of admission control type 3 are Z and Y, respectively. After these configurations, the D-CCAP implements admission control.

Figure 33-44 Admission control diagram

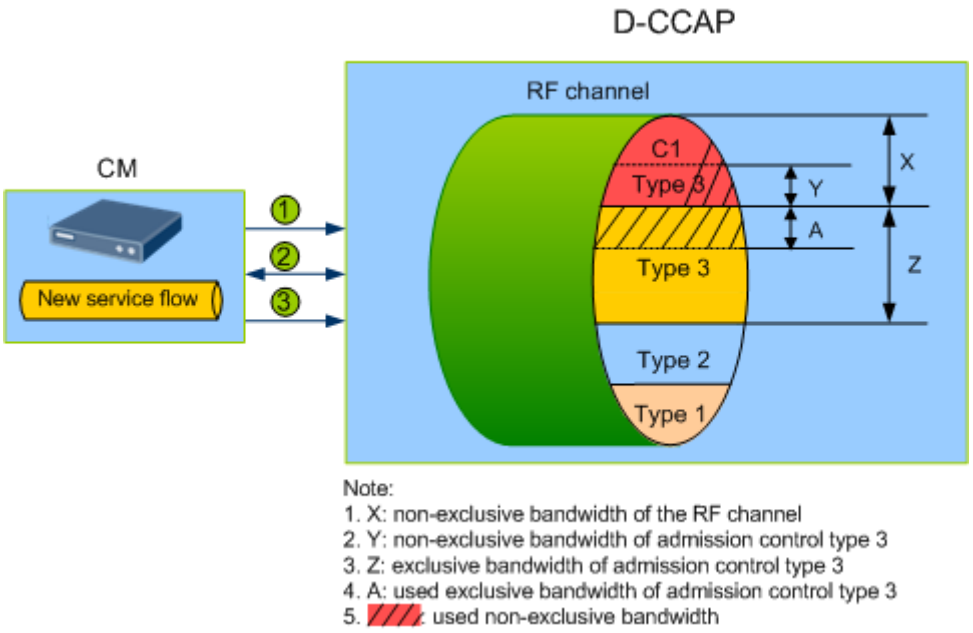


Table 33-19 Admission control process for service flows

No.	Step	Description
1	A CM service flow requests an admission to the D-CCAP.	The D-CCAP controls the following service flow admission requests: • Service flow admission requests initiated by the CM when the CM registers with the D-CCAP (Data services, such as the Internet access

No.	Step	Description
		service, can be enabled on the CM after the CM successfully registers with the D-CCAP.) • Dynamic service flow admission requests, such as for PacketCable calls
2	The D-CCAP matches service flow characteristics to an admission control type. As shown in Figure 33-44, the service flow admission request matches admission control type 3.	The D-CCAP supports eight admission control types. A service flow that has multiple characteristics can match multiple admission control types. When a service flow requests an admission to the D-CCAP, the D-CCAP matches the service flow with the eight admission control types in configuration sequence. The first matched admission control type is used as the admission control type of this service flow. If no admission control type can be matched, the D-CCAP uses the eighth admission control type in configuration sequence for this service flow by default.
3	The D-CCAP checks whether the RF channel bandwidth meets bandwidth requirements of the service flow requesting an admission based on the matched admission control type.	• If the remaining RF channel bandwidth meets the admission requirements, the D-CCAP admits the service flow. • If the remaining RF channel bandwidth does not meet the admission requirements, the D-CCAP performs different operations on the service flow according to the service flow type. – For a CM registration service flow, the D-CCAP searches all downstream and upstream channels for the one that meets the bandwidth requirements. If no channel meets the bandwidth requirements, the D-CCAP rejects the CM registration request to protect the services of online CMs from being affected. – For a dynamic service flow, the D-CCAP rejects the creation request. For example, the voice service of a user fails to create. For instructions about how to check whether the remaining RF channel bandwidth meets the admission requirements, see Bandwidth Determination for Admission Control.

Bandwidth Determination for Admission Control

The D-CCAP implements bandwidth-based admission control for service flows based on used RF channel bandwidths as well as exclusive and non-exclusive bandwidths configured for an admission control type. Determining the bandwidth allocated to admission control involves the following parameters:

- Bandwidth used by the service flows of an admission control type: **Actual**

- Exclusive bandwidth configured for the service flows of an admission control type: **Exclusive**
- Non-exclusive bandwidth configured for the service flows of an admission control type: **Non-exclusive**
- Non-exclusive bandwidth used by the service flows of an admission control type: **Actual-non-exclusive**
- Non-exclusive bandwidth unused by an RF channel: **RF-non-exclusive**
- Minimum assured bandwidth of the service flow requesting an admission: **Require**

In the preceding parameters, **Exclusive** and **Non-exclusive** values are configured and other parameter values are calculated by the D-CCAP. In addition, **Exclusive** and **Non-exclusive** are optional parameters. Table 33-20 lists admission requirements for service flows.

Table 33-20 Admission requirements for service flows

Whether Parameter Exclusive Has Been Configured	Whether Parameter Non-exclusive Has Been Configured	Admission Requirements
Yes	No	• The D-CCAP admits a service flow if the exclusive bandwidth meets the following requirements: **Exclusive** ≥ **Require** + **Actual** • The D-CCAP admits a service flow if the exclusive bandwidth does not meet the admission requirements but the sum of exclusive and non-exclusive bandwidths meets the following requirements: **Exclusive** + **RF-non-exclusive** ≥ **Require** + **Actual**
Yes	Yes	• The D-CCAP admits a service flow if the exclusive bandwidth meets the following requirements: **Exclusive** ≥ **Require** + **Actual** • The D-CCAP admits a service flow if the exclusive bandwidth does not meet the admission requirements but the sum of exclusive and non-exclusive bandwidths meets all of the following requirements: 1. **Non-exclusive** ⩾ **Actual-non-exclusive** 2. **RF-non-exclusive** ⩾ **Non-exclusive** - **Actual-non-exclusive** 3. **Exclusive** + **Non-exclusive** ⩾ **Require** + **Actual**
No	Yes	The D-CCAP admits a service flow if the non-exclusive bandwidth meets all of the following requirements: 1. **Non-exclusive** ≥ **Actual-non-exclusive** 2. **RF-non-exclusive** ≥ **Non-exclusive** - **Actual-non-exclusive** 3. **Non-exclusive** ≥ **Require** + **Actual**

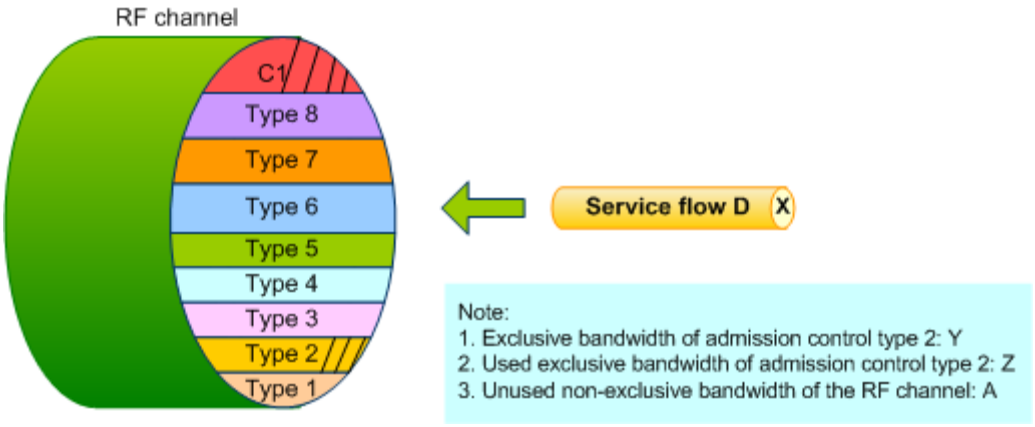
Whether Parameter Exclusive Has Been Configured	Whether Parameter Non-exclusive Has Been Configured	Admission Requirements
No	No	The D-CCAP admits a service flow if the following requirements are met: **RF-non-exclusive** ≥ **Require**

The following provides an example to describe bandwidth determination for admission control if parameter **Exclusive** has been configured and parameter **Non-exclusive** has not been configured. As shown in Figure 33-45,

- Type 1 through type 8 are eight admission control types of the RF channel.
- Service flow D is requesting an admission. The characteristics of service flow D match admission control type 2.
- C1 is the non-exclusive bandwidth of the RF channel.
- X is the minimum assured bandwidth of service flow D.

Service flow D is admitted if the following requirements are met: X ≤ Y - Z + A

Figure 33-45 Admission control example



Bandwidth Exceeding

Bandwidth exceeding is a situation in which the sum of the bandwidths for all service flows of an admission control type exceeds the exclusive bandwidth configured for the service flows of the admission control type. The situation may occur due to operations performed, such as configuring a bandwidth threshold again for downstream and upstream channels, mapping a service flow type, or modifying downstream and upstream channel parameters.

The following provides an example:

- The bandwidth thresholds for the service flows of admission control types 1 and 2 are both 40%.
- The actual bandwidths used by the service flows of admission control types 1 and 2 are 60% and 30%, respectively.

If the bandwidth thresholds for the service flows of admission control types 1 and 2 are changed to 20% and 60%, respectively, the maximum bandwidth for the service flows of admission control type 1 is 40% and the actual used bandwidth is 60%. That is, bandwidth exceeding occurs on the service flows of admission control type 1. When a service flow of admission control type 2 requests an admission to the D-CCAP, the D-CCAP admits the service flow if the admission requirements are met. Then, the D-CCAP schedules queues based on only the priority of Ethernet packets to ensure the bandwidth allocation for the service flows of admission control type 1.

You are advised to plan and configure admission control types before changing the bandwidth thresholds for service flows to prevent bandwidth exceeding.

33.9.4 Configuring Admission Control

This section describes how to configure the admission control feature on the MA5600T/MA5603T/MA5608T.

Procedure

Run the **cable admission-control type** command to name an admission control type.

Step 1 Run the **cable admission-control mapping add** command to configure mapping for the admission control type.

Step 2 (Optional) Run the **cable admission-control bandwidth** command to configure the exclusive bandwidth threshold and non-exclusive bandwidth threshold for upstream and downstream channels.

Step 3 Run the **cable admission-control** command to enable the admission control function.

This function includes the admission control function for CM registration and the admission control function for dynamic service flow creation.

----End

Example

The following is an example of the configurations used to enable the admission control feature:

- A 10% exclusive bandwidth is configured for emergency calls.
- A 5% non-exclusive bandwidth is configured for emergency calls.
- The service flows meeting emergency call requirements have two characteristics:
  - The priority of PacketCable sessions is high.
  - The priority of the **SessionClassID** field in COPS is 7 for PacketCable Multimedia service flows.
- Service flows with the two preceding characteristics are mapped to admission control type 0.
- Admission control type 0 is named **emergency-call**.

```
huawei(config)#cable admission-control type 0 name emergency-call
huawei(config)#cable admission-control mapping add type 0 sub-type packetcable 1dotx
high-priority
huawei(config)#cable admission-control mapping add type 0 sub-type packetcable multim
edia priority 7
```

```
huawei(config)#interface cable 1/1/0
huawei(config-if-cable-1/1/0)#cable admission-control bandwidth upstream type 0 exclusive 10 non-exclusive 5
huawei(config-if-cable-1/1/0)#cable admission-control bandwidth downstream type 0 exclusive 10 non-exclusive 5
huawei(config-if-cable-1/1/0)#quit
huawei(config)#cable admission-control dynamic-service enable
```

The following is an example of the configurations used to enable the admission control feature:

- A 10% exclusive bandwidth is configured for CM registration service flows.
- The service flows meeting CM registration service flow requirements have two characteristics:
  - The service class name of the service flows is **huawei**.
  - The upstream service flows are UGS service flows.
- Service flows with the two preceding characteristics are mapped to admission control type 1.
- Admission control type 1 is named **cm1**.

```
huawei(config)#cable admission-control type 1 name cm1
huawei(config)#cable admission-control mapping add type 1 sub-type service-class-name huawei
huawei(config)#cable admission-control mapping add type 1 sub-type scheduling-type ugs
huawei(config)#interface cable 1/1/0
huawei(config-if-cable-1/1/0)#cable admission-control bandwidth upstream type 1 exclusive 10
huawei(config-if-cable-1/1/0)#cable admission-control bandwidth downstream type 1 exclusive 10
huawei(config-if-cable-1/1/0)#quit
huawei(config)#cable admission-control cm-registration enable
```

33.9.5 Standards and Protocols Compliance

The standards and protocols that the admission control feature complies with are as follows:

- CM-TR-OSSIv3.0-CM-V01-08092
- CM-TR-MGMTv3.0-DIFF-V01-071228
- CM-SP-SECv3.0-I13-100611
- CM-SP-PHYv3.0-I09-101008
- CM-SP-DRFI-I11-110210
- CM-SP-OSSIv3.0-I14-110210
- CM-SP-MULPIv3.0-I15-110210
- C-DOCSISv2.1

33.10 QoS Adjustment

What Is QoS Adjustment

Definition

QoS adjustment is a traffic management mechanism provided by the D-CCAP. It prevents a small number of users from occupying a large amount of bandwidth.

The D-CCAP implements QoS adjustment by continuously monitoring users' bandwidth usages. If a user occupies a large amount of bandwidth for a long period of time, the D-CCAP reduces the user's service level agreement (SLA) to lower the user's bandwidth and service flow priority. This QoS adjustment allows the D-CCAP to restrict users from occupying large amounts of bandwidth for long periods of time.

Benefits

Home users subscribe to monthly bandwidth packages offered by broadcast and television companies. If some home users continuously consume a large amount of bandwidth, the other home users are unable to normally use the network. QoS adjustment is used to help resolve this problem, and provides the following benefits to carriers:

- Prevents a small number of users from occupying a large amount of bandwidths, improving user experience.
- Dynamically limits traffic and bandwidth of certain users, allowing carriers to implement fine-grained traffic management not possible with standard monthly packages.

33.10.2 Basic Concepts

QoS adjustment involves the concepts listed in the following table.

Table 33-21 Basic QoS adjustment concepts

Concept	Description
Monitoring mode	QoS adjustment monitors average bandwidth or service traffic in peak monitoring or off-peak monitoring mode.
Peak monitoring	In this mode, the D-CCAP monitors average bandwidth or service traffic during peak hours and sets different monitoring durations and rates for peak hours than for other monitoring hours. Therefore, this mode provides flexible monitoring.
Off-peak monitoring	In this mode, the MA5633 monitors average bandwidth and service traffic during off-peak hours with the same monitoring duration and rate.
Sampling period	A certain number of sampling points are set in a monitoring period to periodically sample the rates of specified service flows. The duration between two sampling points is called a sampling interval or a sampling period.
Monitoring window	The D-CCAP determines whether the service traffic exceeds the upper threshold in this window. If the service traffic exceeds the upper threshold, the D-CCAP performs QoS adjustment. Otherwise,

Concept	Description
	the D-CCAP proceeds to the next monitoring window.
Adjustment period	QoS adjustment is performed if the rate of a service flow exceeds a preset threshold when the duration of a monitoring window ends. The QoS adjustment profile is bound to the service flow during the entire adjustment period. After the adjustment period ends, the basic QoS profile is re-bound to the service flow. Then, the D-CCAP starts another round of monitoring.

33.10.3 QoS Adjustment Process

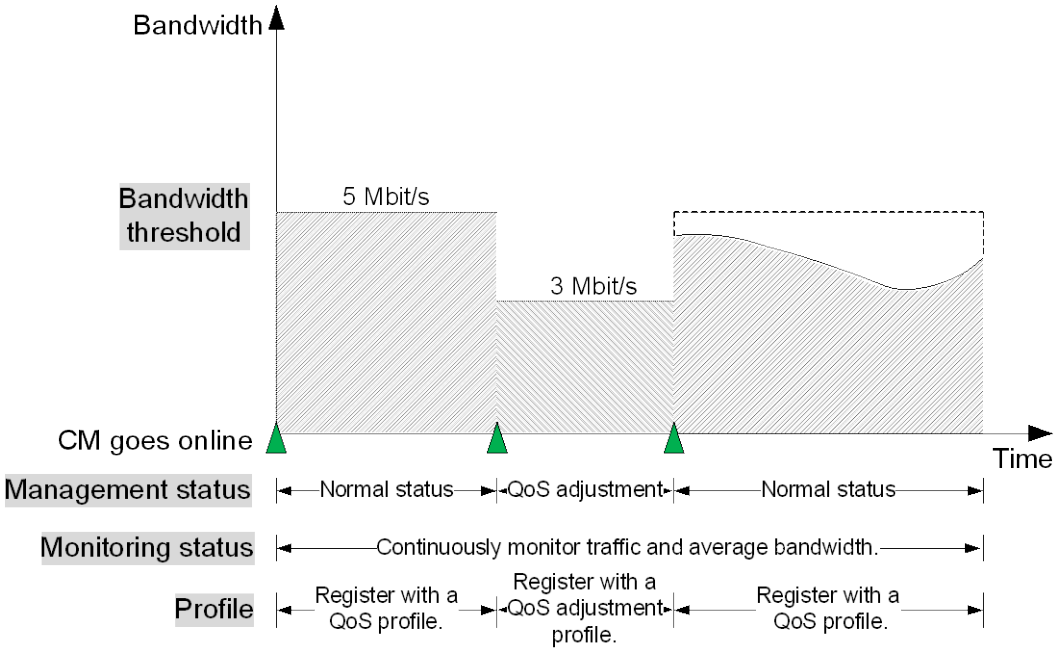
In the QoS adjustment process, a basic QoS profile (the business bandwidth package profile for home users) and a QoS adjustment profile are configured on a distributed converged cable access Platform (D-CCAP). The QoS adjustment profile contains the average bandwidth threshold, average traffic threshold, and QoS adjustment bandwidth.

In the following example, the bandwidth provided by the basic QoS profile is 5 Mbit/s, and the average bandwidth threshold for the QoS adjustment profile is 3 Mbit/s.

When a cable modem (CM) goes on line, it registers with a basic QoS profile and uses the subscribed 5 Mbit/s bandwidth package. The D-CCAP continuously monitors the used traffic or average bandwidth of the user (the CM) and determines whether the used traffic or average bandwidth exceeds the thresholds. If the D-CCAP detects that a threshold has been exceeded, it starts QoS adjustment. Specifically, the D-CCAP registers with a QoS adjustment profile and reduces the user's bandwidth to 3 Mbit/s. After the QoS adjustment period expires, the CM registers with the basic QoS profile again and has a bandwidth of 5 Mbit/s.

The following figure shows the QoS adjustment process for this example. In the "Normal status" phase, the CM registers with a basic QoS profile (the applied business package profile). In the "QoS adjustment period" phase, the CM registers with a QoS adjustment profile.

Figure 33-46 Process of QoS adjustment

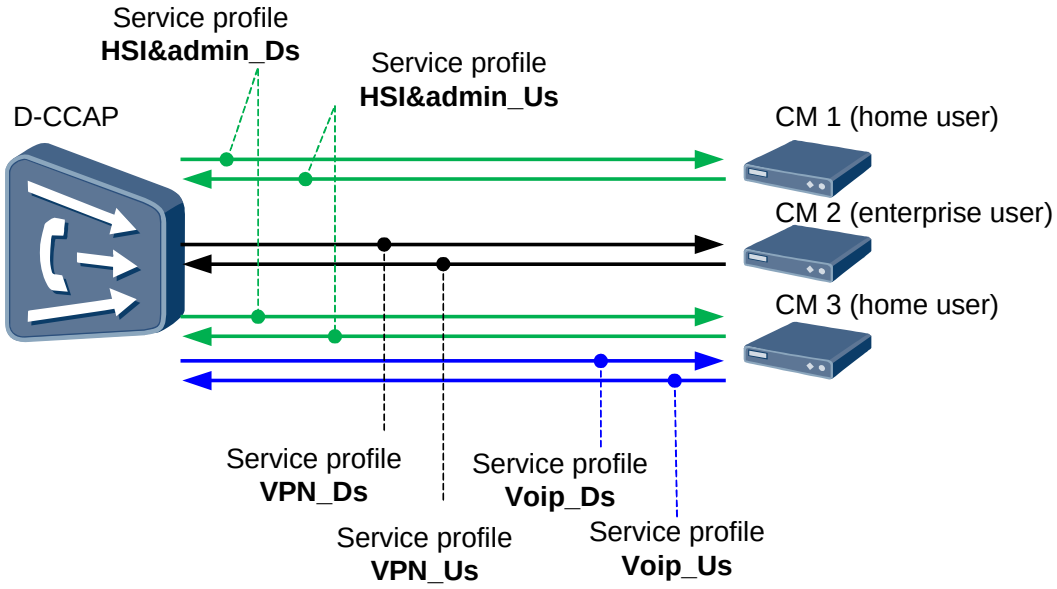


33.10.4 Configuring QoS Adjustment on Service Flows

QoS adjustment is used for data services of home users. The voice and private line services have the highest priorities and also have high quality of service (QoS) requirements. Therefore, QoS adjustment is not performed on these two services.

As shown in Figure 33-47, a distributed converged cable access Platform (D-CCAP), is connected to 3 cable modems (CMs). CM 1 and CM 3 are home users and CM2 is an enterprise user.

Figure 33-47 Configuring QoS adjustment on service flows



- CM 1 uses one pair of service flows to carry the HSI and management services.
- CM 3 uses two pairs of service flows. One pair (shown in green) carries the HSI and management services and the other pair (shown in blue) carries the voice service.

In this scenario, QoS adjustment needs to be implemented on the HSI and management services of home users (CM 1 and CM 3) but not on the voice service of home users (CM 3) and all services of the enterprise user (CM 2).

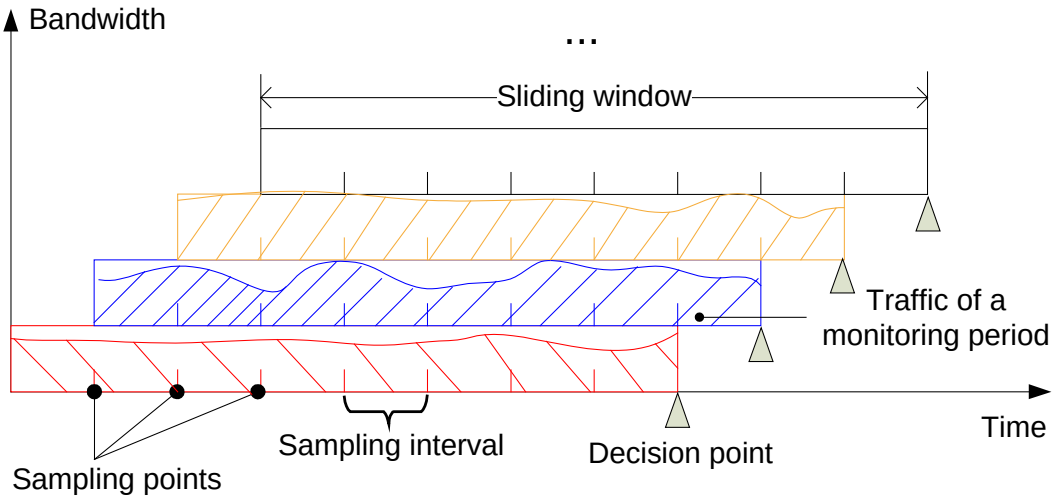
To meet the preceding requirements, perform the following operations:

- Configure QoS adjustment on the service profiles **HSI&admin_Ds** and **HSI&admin_Us**.
- Configure the average bandwidth, traffic thresholds, and QoS adjustment profile for the service profiles **HSI&admin_Ds** and **HSI&admin_Us**.

33.10.5 Sampling, Monitoring, and Decision Making

To implement quality of service (QoS) adjustment, the D-CCAP periodically collects samples from, monitors, and makes decisions for specified service flows, as shown in Figure 33-48.

Figure 33-48 Sampling, monitoring, and decision making



Sampling

During a monitoring period, the D-CCAP periodically samples the rates of specified service flows. The duration between two sampling points is called a sampling interval or a sampling period. The default sampling period is 20 minutes.

The D-CCAP determines whether the service flows require QoS adjustment based on the results of multiple sampling periods.

Monitoring

The D-CCAP monitors service flows through a monitoring window. The duration of monitoring window is a monitoring period. When a monitoring period expires, the D-CCAP monitors whether the traffic or average bandwidth of a service flow exceeds the preset threshold and then determines whether QoS adjustment is required.

- If the traffic or average bandwidth exceeds the preset threshold, the D-CCAP stops monitoring and starts QoS adjustment.
- If the traffic or average bandwidth does not exceed the preset threshold, the D-CCAP proceeds to the next monitoring window.

The default monitoring period is 3 hours. A monitoring period can include multiple monitoring windows, as shown in Figure 33-48.

The D-CCAP supports phase-based monitoring, which is classified into three modes: peak monitoring, off-peak monitoring, and weekend monitoring.

- Peak monitoring: The D-CCAP monitors traffic in peak hour phases (a maximum of 2 peak hour phases can be set) during the monitoring period and sets different monitoring duration and rates for this peak hour phase from other monitoring hours, as shown in Figure 33-49.
- Off-peak monitoring: The D-CCAP continuously monitors traffic for all hours during the monitoring period with the same monitoring duration and rate, as shown in Figure 33-50.

- Weekend monitoring: The D-CCAP monitors traffic only on Saturday and Sunday in either peak or off-peak monitoring mode. Figure 33-51 shows weekend monitoring in peak monitoring mode.

Figure 33-49 Peak monitoring

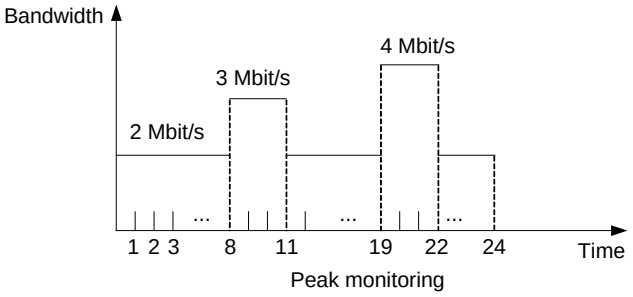


Figure 33-50 Off-peak monitoring

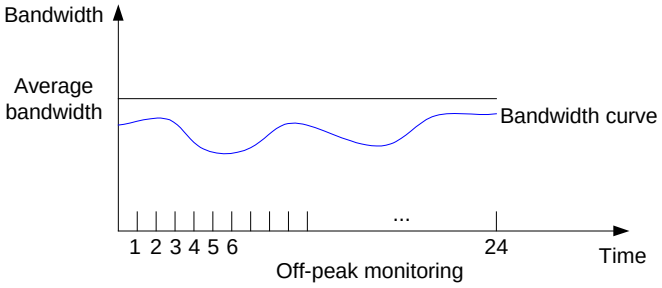
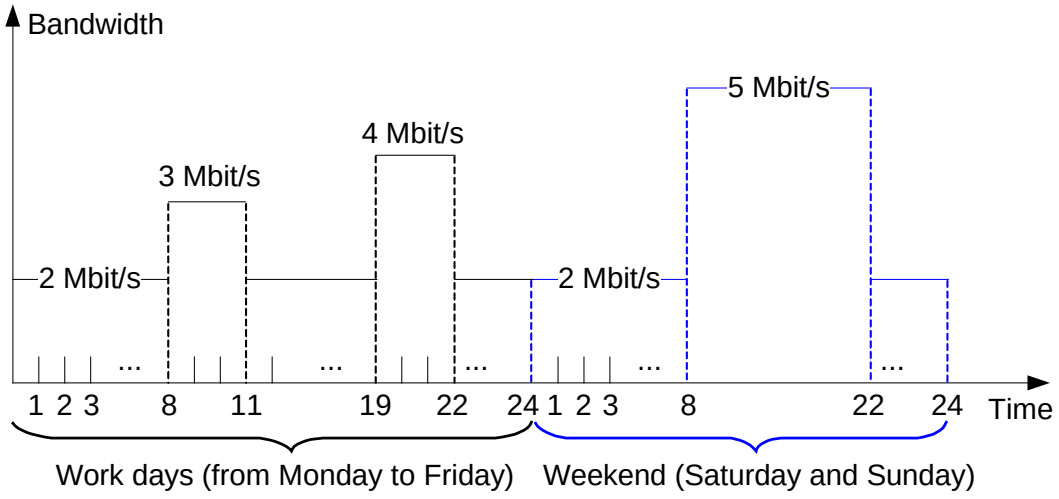


Figure 33-51 Weekend monitoring in peak monitoring mode



Decision Making

When a monitoring window expires, the D-CCAP decides whether to start QoS adjustment based on whether the traffic or average rate exceeds the preset threshold during the monitoring period.

Decision making requires two parameters:

- Traffic: indicates whether the total traffic during a monitoring period exceeds the preset threshold. This parameter is valid only in off-peak monitoring mode and is optional.
- Average bandwidth: indicates whether the average bandwidth during a monitoring period exceeds the preset threshold. This parameter is mandatory.

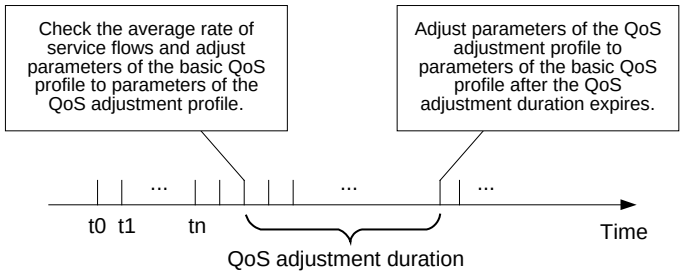
If both parameters are configured in off-peak monitoring mode, the D-CCAP starts QoS adjustment if either parameter exceeds the preset threshold. In peak monitoring mode, the D-CCAP starts QoS adjustment if the average bandwidth exceeds the preset threshold.

33.10.6 QoS Adjustment Principles

QoS Adjustment Principles

If the D-CCAP detects a threshold-exceeding event indicating that the traffic or average bandwidth of a CM exceeds the preset threshold, it binds upstream and downstream service flows to a QoS adjustment profile to lower the bandwidth or priorities of the service users, as shown in the following figure.

Figure 33-52 QoS adjustment principles



1. If the D-CCAP detects that the rate of a service flow exceeds the preset threshold, the D-CCAP adjusts the QoS parameter settings of this service flow to those in the QoS adjustment profile.
2. The D-CCAP starts the QoS adjustment timer and changes the service flow status to QoS adjustment.
3. When the timer times out, the D-CCAP adjusts the QoS parameter settings of the service flow to those recorded in the configuration file and changes the service flow status to normal.
4. During QoS adjustment, the D-CCAP samples the traffic of services but does not determine whether the service flow rate exceeds the preset threshold.

Exiting QoS Adjustment

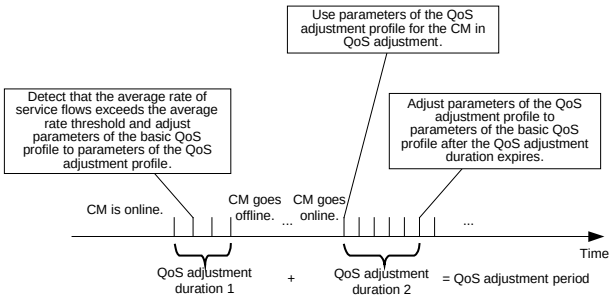
The D-CCAP supports two modes for exiting QoS adjustment:

- Timeout exiting: In this mode, the QoS adjustment period is configurable. A CM exits QoS adjustment when the period expires. This mode applies to fixed QoS adjustment periods. The default QoS adjustment period is 1 day.
- Forcibly exiting: In this mode, a CM forcibly exits QoS adjustment by running the **cable qos adjust cancel cm** command before the QoS adjustment period expires.

Processing When a CM Resets During QoS Adjustment

Figure 33-53 shows the actions performed by the D-CCAP when a CM resets during a QoS adjustment period.

Figure 33-53 Processing when a CM resets during QoS adjustment period



When a CM resets during a QoS adjustment period, the D-CCAP performs the following actions:

1. Records the used QoS adjustment duration (QoS adjustment duration 1 in Figure 33-53).
2. Determines whether the QoS adjustment period for this CM expires when the CM goes online again. If the QoS adjustment period does not expire, the D-CCAP binds the parameter settings of the QoS adjustment profile to the CM and sets the timer to the remaining QoS adjustment duration.
3. Changes QoS parameter settings of the CM to those in the basic QoS profile after the remaining QoS adjustment duration expires.

In addition, the D-CCAP allows users to configure whether to resume QoS adjustment after the CM resets.

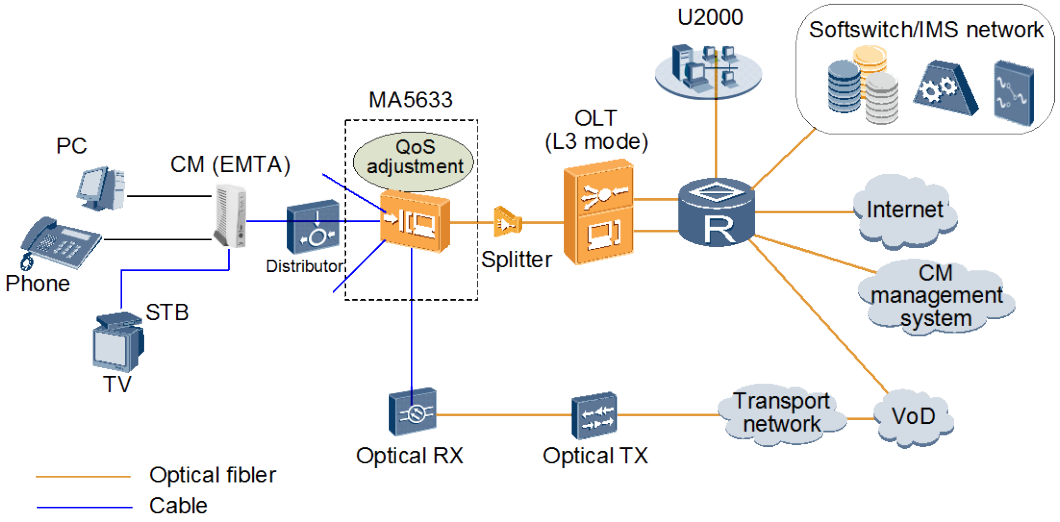
33.10.7 Networking Applications

QoS adjustment can be used in both D-CCAP standalone NE management networking and D-CCAP centralized management networking.

Standalone NE Management Networking

Figure 33-54 shows QoS adjustment in standalone NE management networking.

Figure 33-54 QoS adjustment in standalone NE management networking

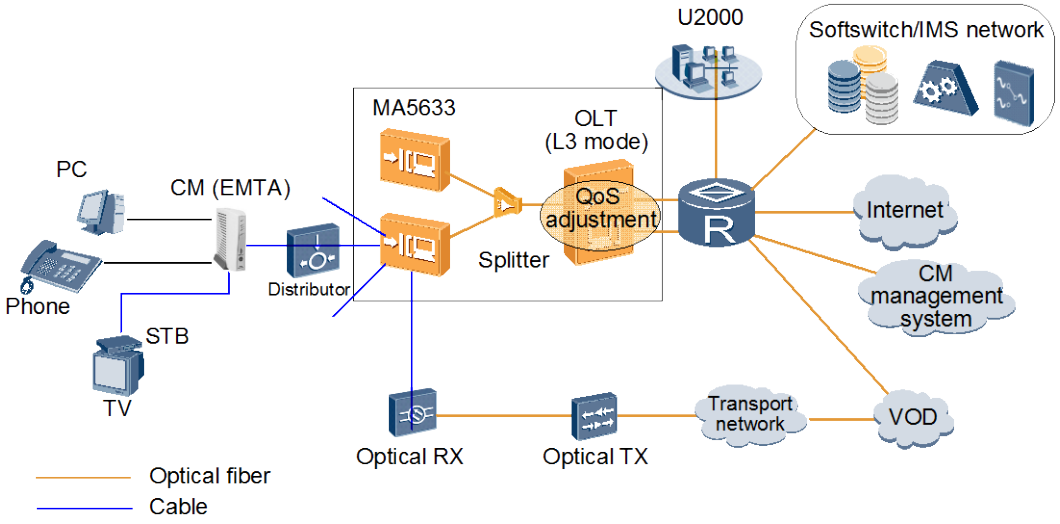


- The CMC is connected upstream to the optical line terminal (OLT) through a PON port or a GE port. The OLT is connected upstream to the IP network through a router or Layer 3 switch.
- The OLT working in Layer 3 forwarding mode and the CMC use standalone NE management.
- QoS adjustment is enabled on the CMC to monitor traffic of Internet access and data services for home users. The CMC identifies users occupying large amounts of bandwidth and implements QoS adjustment accordingly.

CMTS Centralized Management Networking

Figure 33-55 shows QoS adjustment in centralized management networking.

Figure 33-55 QoS adjustment in centralized management networking



- The CMC is connected upstream to the OLT through a PON port or a GE port. The OLT is connected upstream to the IP network through a router or Layer 3 switch.
- The OLT working in Layer 3 forwarding mode and the CMC use centralized management networking.

NOTE

In this mode, the CMC serves as an extended subrack of the OLT. All parameters are configured on the OLT and issued to the CMC.

- QoS adjustment is enabled on the OLT to monitor traffic of Internet access and data services for home users, and is deployed on the CMC. The CMC identifies users occupying large amounts of bandwidth and implements QoS adjustment accordingly.

33.10.8 Configuring QoS Adjustment

Prerequisites

Services have been configured on an optical line terminal (OLT) and a cable modem (CM). Both the OLT and CM are functional.

Context

QoS adjustment can be used in both D-CCAP independent networking and D-CCAP centralized management networking. These two networking scenarios have the following differences:

- In independent management networking, QoS adjustment is configured on the MA5633.
- In centralized management networking, QoS adjustment is configured on the OLT.

Procedure

Configure the QoS profile and QoS adjustment profile in both the upstream and downstream directions.

Run the **cable service-class** command to configure the QoS profile and QoS adjustment profile in both the upstream and downstream directions.

Step 1 Configure the upstream and downstream QoS adjustment.

1. Configure the monitoring mode of QoS adjustment.

   Run the **monitoring-mode** command to configure the monitoring mode of QoS adjustment.

   - The monitoring mode supports peak monitoring and non-peak monitoring.
   - The monitoring mode can be separately set for the upstream and downstream directions.

2. (Optional) Configure the peak monitoring parameters of QoS adjustment.

   To enable peak monitoring, run the **peak-monitoring** command to configure the peak monitoring parameters of QoS adjustment.

3. (Optional) Configure the non-peak monitoring parameters of QoS adjustment.

   To enable non-peak monitoring, run the **nopeak-monitoring** command to configure non-peak monitoring parameters of QoS adjustment.

4. Configure the QoS adjustment period.

Run the **adjust-period** command to set the QoS profile to QoS adjustment profile (for QoS adjustment) and configure the QoS adjustment duration, QoS adjustment end time, and whether further monitoring is required after QoS adjustment stops.

5. (Optional) Configure the traffic threshold for QoS adjustment.

   QoS adjustment is determined by thresholds of the average rate and traffic for a service stream. The traffic threshold is valid only in non-peak monitoring mode. In non-peak monitoring mode, both the thresholds of the average rate and traffic can be configured. QoS adjustment starts if either of the average rate and traffic for a service stream exceeds the preset threshold.

   Run the **rule active-threshold** command to configure the traffic threshold for QoS adjustment.

Step 2 Activate upstream and downstream QoS adjustment.

Run the **active** command to activate upstream and downstream QoS adjustment.

Step 3 Query QoS adjustment information.

- Run the **display cable qos adjust-rule** command to query details of QoS adjustment.
- Run the **display cable qos-adjust** command to query information about QoS adjustment of a specified cable modem (CM).

----End

Example

The following is an example of the configurations used to globally enable QoS adjustment on the MA5633 for a CM in an independent management networking scenario. These configurations are used to prevent a small number of users from occupying a large amount of bandwidth. Configure the parameters as follows:

- Set upstream QoS profile to **up-qos**, downstream QoS profile to **ds-qos**, maximum burst bytes to 3044, and maximum stable rate to 2 Mbit/s (2097152 bit/s)
- Set upstream QoS profile to **up-adjust-qos**, downstream QoS profile to **ds-adjust-qos**, maximum burst bytes to 2500, and maximum stable rate to 1 Mbit/s (1048576 bit/s)
- Configure the parameters of downstream QoS adjustment **ds-rule** as follows:
  - Set bound QoS profile to **ds-qos** and QoS adjustment profile to **ds-adjust-qos**
  - For peak hour monitoring: Set peak hour 1 to 12:00 with a 60-minute monitoring window; set peak hour 2 to 19:00 with a 180-minute monitoring window; set non-peak hours to a 60-minute monitoring window.
  - Set average upstream bandwidth of peak hour 1 to 2000 kbit/s, average upstream bandwidth of peak hour 2 to 4000 kbit/s, sampling period to 15 minutes, and average upstream bandwidth of non-peak hours to 500 kbit/s.
  - Set QoS adjustment period to 30 minutes and adjustment resetting time to 23:00. No further monitoring is performed after resetting.
- Configure the parameters of upstream QoS adjustment up-rule as follows:
  - Set bound QoS profile to **up-qos** and QoS adjustment profile to **up-adjust-qos**.
  - For peak hour monitoring: Set peak hour 1 to 12:00 with a 60-minute monitoring window; set peak hour 2 to 19:00 with a 180-minute monitoring window; set non-peak hours to a 60-minute monitoring window.

- Set average upstream bandwidth of peak hour 1 to 1500 kbit/s, average upstream bandwidth of peak hour 2 to 3000 kbit/s, sampling period to 15 minutes, and average upstream bandwidth of non-peak hours to 500 kbit/s.
- Set QoS adjustment period to 30 minutes and adjustment resetting time to 23:00. No further monitoring is performed after resetting.

The configuration process is shown as follows:

```
//Configure the upstream QoS profile
huawei(config)#cable service-class name up-qos upstream max-burst 3044 max-rate 2097152
//Configure the downstream QoS profile
huawei(config)#cable service-class name ds-qos downstream max-burst 3044 max-rate
2097152
//Configure the upstream QoS adjustment profile
huawei(config)#cable service-class name up-adjust-qos upstream max-burst 2500 max-rate
1048576
//Configure the downstream QoS adjustment profile
huawei(config)#cable service-class name ds-adjust-qos downstream max-burst 2500
max-rate 1048576

//Configure the downstream QoS adjustment
huawei(config)#cable qos adjust-rule ds-rule
huawei(config-cable-qos-adjust-ds-rule)#monitoring-mode peak
huawei(config-cable-qos-adjust-ds-rule)#service-class basic ds-qos
huawei(config-cable-qos-adjust-ds-rule)#service-class adjust ds-adjust-qos
huawei(config-cable-qos-adjust-ds-rule)#peak-monitoring peak-time1 12:00 window 60
average-rate 2000 peak-time2 19:00
 window 180 average-rate 4000 remaining 60 average-rate 500 sampling-interval 15
downstream
huawei(config-cable-qos-adjust-ds-rule)#adjust-period 30 reset-time 23:00
huawei(config-cable-qos-adjust-ds-rule)#active

//Configure the upstream QoS adjustment
huawei(config)#cable qos adjust-rule up-rule
huawei(config-cable-qos-adjust-up-rule)#monitoring-mode peak
huawei(config-cable-qos-adjust-up-rule)#service-class basic up-qos
huawei(config-cable-qos-adjust-up-rule)#service-class adjust up-adjust-qos
huawei(config-cable-qos-adjust-up-rule)#peak-monitoring peak-time1 12:00 window 60
average-rate 1500 peak-time2 19:00
 window 180 average-rate 3000 remaining 60 average-rate 500 sampling-interval 15
upstream
huawei(config-cable-qos-adjust-up-rule)#adjust-period 30 reset-time 23:00
huawei(config-cable-qos-adjust-up-rule)#active
```

33.11 SAV

The source address verification (SAV) feature enables an MA5600T/MA5603T/MA5608T to verify the source IP addresses of received packets. The SAV feature prevents unauthorized users from accessing the system and ensures that the system properly processes the services of authorized users.

Introduction

If a malicious user forges the IP address of an authorized user and sends a great number of packets to attack a system, the system cannot process the services of authorized users.

The SAV feature enables an MA5600T/MA5603T/MA5608T to bind the authorized IP addresses to corresponding users to implement the following functions:

- Protect carrier networks from being attacked.
- Ensure service security.

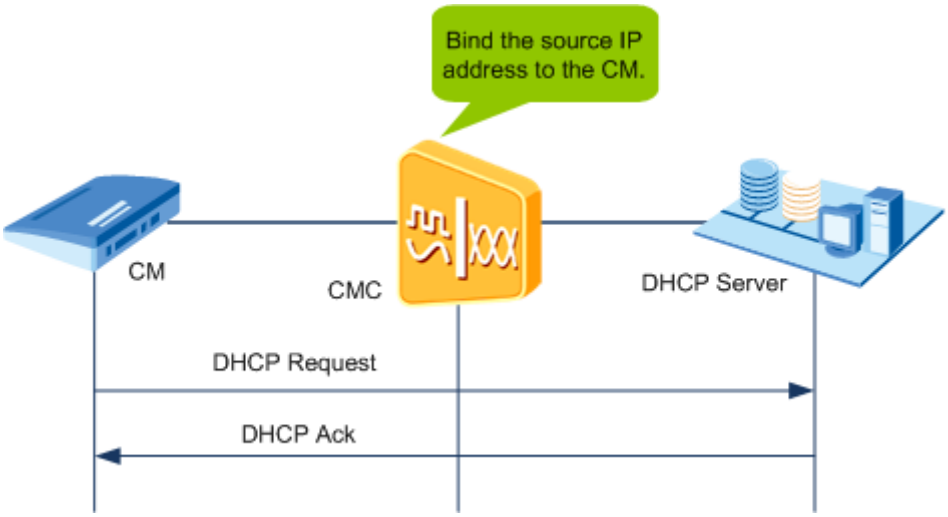
33.11.1 Principles

After the SAV feature is enabled, the MA5600T/MA5603T/MA5608T drops upstream packets from unauthorized source IP addresses.

The IP address of the MA5600T/MA5603T/MA5608T can be obtained from a Dynamic Host Configuration Protocol (DHCP) packet or statically specified. Authorized IP addresses can also be dynamically or statically bound to the SAV feature.

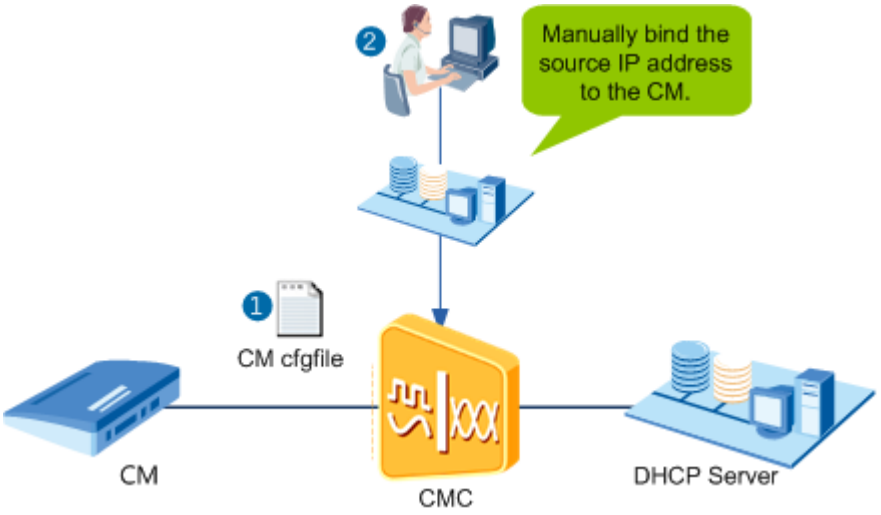
Dynamic SAV IP Address Binding

- The system monitors DHCP online and offline processes for users. When a user goes online, the system dynamically obtains the source IP address of the user and binds the IP address to the cable modem (CM) of the user.
- Only service packets with an authorized IP address bound to a CM can pass through the MA5600T/MA5603T/MA5608T.
- When the user goes offline, the system unbinds the source IP address from the CM.



Static SAV IP Address Binding

Users can configure static SAV IP address binding rules through the network management system (NMS), command line interface (CLI), or CM configuration file. Using these rules, the system binds the source IP address of a user to the CM of the user when the user goes online.



The parameters in the CM configuration file define the range of bound static IP addresses.

The parameters in the CM configuration file for defining the range of bound static IP addresses contain the following type-length-value (TLV) formats:

- SAV prefix group ID encoding

  Packet filtering is configured in the MA5600T/MA5603T/MA5608T. Only packets with IP addresses in a defined network segment can pass through the MA5600T/MA5603T/MA5608T.

  Specifically, an SAV name and its IP address segment are configured in the MA5600T/MA5603T/MA5608T. When a CM starts or the configuration of a CM restores, the configuration file carries only the SAV name. The MA5600T/MA5603T/MA5608T obtains the SAV name from the configuration file, identifies the IP address segment, and issues entries, implementing packet filtering.
- Static SAV prefix encoding

  The range of static IP addresses is defined in upstream service flows of a CM.

  Specifically, users do not need to configure an SAV name or its IP address segment in the MA5600T/MA5603T/MA5608T. When a CM starts or the configuration of a CM restores, the configuration file carries SAV rule IDs and their IP address segments and the MA5600T/MA5603T/MA5608T issues entries, implementing packet filtering.

A valid CM configuration file contains only one SAV prefix group ID encoding TLV or one or multiple static SAV prefix encoding TLVs.

TLV Format

Figure 33-56 shows the SAV TLV format in the CM configuration file.

Figure 33-56 TLV format in the CM configuration file

TLV	Name	
43.7.1	SAV Group Name Subtype	SAV Prefix Group ID Encoding
43.7.2	SAV Static Prefix Rule Subtype	Static SAV Prefix Encoding
43.7.2.1	SAV Static Prefix Address Subtype	
43.7.2.2	SAV Static Prefix Length Subtype	

33.11.2 Configuring SAV

This section describes how to configure the SAV feature to prevent malicious users from attacking the MA5600T/MA5603T/MA5608T or authorized users by forging the IP addresses of authorized users.

Context

The IP address of the MA5600T/MA5603T/MA5608T can be obtained from a DHCP packet or statically configured. Accordingly, authorized IP addresses can be dynamically or statically bound to the CMs of users.

- Dynamic IP address binding: The system monitors DHCP user online and offline processes. When a DHCP user goes online, the system dynamically obtains the source IP address of the user and binds the IP address to the CM of the user.
- Static IP address binding: Users can configure a permitted static IP address network segment using the NMS, CLI, or CM configuration file. When a user goes online, the system binds the source IP address (the configured static IP address network segment) of the user to the CM of the user.

Only service packets with an authorized IP address bound to a CM can pass through the MA5600T/MA5603T/MA5608T.

Procedure

- Configure IPv4 SAV.
  a. Run the **cable source-verify enable** command to enable IPv4 SAV in global config mode and VLAN service profile mode, respectively.
  b. Run the **cable source-verify group** command to configure the permitted static IP address network segment. After the configuration, only packets within this IP address network segment can be transmitted.
- Configure IPv6 SAV.
  a. Run the **cable ipv6 source-verify enable** command to enable IPv6 SAV in global config mode and VLAN service profile mode, respectively.
  b. Run the **cable source-verify group** command to configure the permitted static IPv6 address network segment. After the configuration, only packets within this IPv6 address network segment can be transmitted.

----End

Example

The IPv4 SAV feature needs to be enabled to ensure network security. The following is an example of configurations used to configure the SAV feature:

- SAV group name: **sav1**
- Permitted network segment 1: **10.10.10.10/24**
- Permitted network segment 2: **10.20.20.20/24**

In addition, configure the SAV name in the CM configuration file to enable the MA5600T/MA5603T/MA5608T to parse the SAV name in the configuration file, identify the IP address network segments, and issue entries when a CM goes online to filter IP address network segments. The following is an example of configurations used to enable users in the permitted network segments to access the network:

```
huawei(config)#cable source-verify enable
huawei(config)#cable source-verify group sav1 rule 1 ip 10.10.10.10 24
huawei(config)#cable source-verify group sav1 rule 2 ip 10.20.20.20 24
```

The IPv6 SAV feature needs to be enabled to ensure network security. The following is an example of configurations used to configure the SAV feature:

- SAV group name: **sav2**
- Permitted network segment 1: **2001:DB8::1/32**

```
huawei(config)#cable ipv6 source-verify enable
huawei(config)#vlan service-profile profile-id 2
huawei(config-vlan-srvprof-2)#cable ipv6 source-verify enable
huawei(config)#cable source-verify group sav2 rule 1 ip 2001:DB8::1 32
```

33.11.3 SAV Standards and Protocols Compliance

- CM-SP-MULPIv3_0-I09-090121.pdf
- CM-SP-OSSIv3.0-I17-111221.pdf
- CM-SP-SECv3.0-I09-090121.pdf

33.12 Validity Check for a CM

The MA5600T/MA5603T/MA5608T supports the validity check for a cable modem (CM) feature. This feature prevents theft of network resources by an unauthorized CM.

Introduction

Common Internet security threats include the following:

- Unauthorized use: Resources are used without authorization. For example, attackers gain access to a computer system and use resources by guessing a user's account name and password.
- Information theft: Attackers do not invade the target system, but instead sniff the system to steal important data or information.

BPI+ authentication encrypts service flows between the MA5600T/MA5603T/MA5608T and CMs to implement data encryption on a hybrid fiber coaxial (HFC) network. In addition, BPI+ authentication provides authentication parameters and service encryption cipher keys (CKs) to ensure service security between the MA5600T/MA5603T/MA5608T and CMs.

BPI+ authentication enhances service security in the following two aspects:

- Authenticates the CM digital certificate using X.509 authentication and therefore prevents unauthorized CMs from going online.
- Prevents unauthorized CMs from intercepting service data.

33.12.1 Principles

BPI+ and EAE

According to the baseline privacy interface (BPI) standard defined in data over cable service interface specification (DOCSIS), the CK management protocol cannot authenticate a CM and therefore the BPI feature cannot meet service protection requirements. BPI+ authentication resolves this issue. Specifically, the BPI+ feature enhances service protection by supporting CM-based X.509 digital certificate authentication.

The D-CCAP provisions services for and authenticates a CM by the MAC address of the CM. Therefore, the security of the CM MAC address must be ensured, and the network must be protected from unauthorized CMs that may access the network by duplicating the MAC address of an authorized CM.

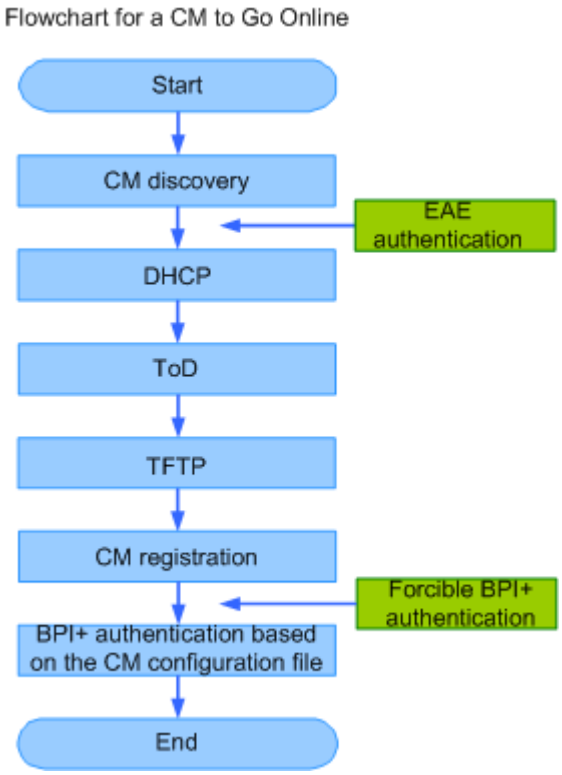
According to the BPI+ feature, a CM must carry an X.509 digital certificate. The X.509 digital certificate contains the valid MAC address of the CM and uses the manufacturer certificate for the digital signature to verify the validity of the CM. When a CM attempts to access a network, the D-CCAP performs X.509 authentication for the digital certificate reported by the CM. This operation is performed to check whether the MAC address in the digital certificate is the same as that of the CM. If the MAC addresses are the same, the D-CCAP considers the CM to be authorized and allows it to access the network. Otherwise, the D-CCAP does not allow the CM to access the network. The CM digital certificate uses the signature of the manufacturer certificate, which is difficult to forge, enhancing the security of the D-CCAP.

BPI+ authentication is performed at the CM registration stage. The D-CCAP starts BPI+ authentication upon receiving a CM authentication message. D-CCAP supports forcible BPI+ authentication. This feature allows the D-CCAP to forcibly perform BPI+ authentication for all CMs or a user connected to a specified port.

Early authentication and encryption (EAE) is a secure authentication mode introduced in DOCSIS 3.0. EAE is basically the same as the BPI+ authentication excepting that the authentication is performed before the DHCP stage (the BPI+ authentication is performed at the CM registration stage). EAE authentication messages are the same as BPI+ authentication messages, which ensures the encryption of all packets from the DHCP stage. This authentication enhances system security.

Figure 33-57 shows the stages at which EAE authentication, forcible BPI+ authentication, and common BPI+ authentication are performed during the CM online process. Basically, the three types of authentication are BPI+ authentication. They vary depending on the authentication stage and authentication mode (forcible or not). If the authentication is performed early during the CM online process, packets are encrypted early, thereby enhancing system security.

Figure 33-57 Stages at which EAE authentication, forcible BPI+ authentication, and common BPI+ authentication are performed



X.509 Digital Certificate

According to the baseline privacy interface plus (BPI+) standard, a CM must carry an X.509 digital certificate to achieve system security.

The X.509 digital certificate includes information about the applicant and CA certificate issuance. All digital certificates issued by the authentication center comply with X.509v3. The International Telecommunication Union (ITU) and X.509v3 define the format of the digital certificate.

Classification

According to the X.509 authentication process defined in data over cable service interface specification (DOCSIS), the digital certificate is managed in Layer 3 management mode.

1. The Cable Television Laboratories (CableLabs) issues and maintains root certificate authority (CA) in a unified manner.
2. Manufacturers apply to CableLabs for manufacturer CA certificates.
3. The manufacturers use their manufacturer CA certificates to issue certificates for the CMs they manufacture.

Table 33-22 describes the certificates used for verify CM validity.

Table 33-22 Certificates used to verify CM validity

Certificate Type	Purpose	Signing Mode	Remarks
Root CA certificate	• Verifies the validity of a manufacturer CA certificate. • Self verifies its validity.	The public key certificate of a root CA certificate is signed using the root CA certificate itself.	The MA5600T/MA5603T/ MA5608T supports root CA certificate importing to verify the validity of a manufacturer CA certificate.
Manufacturer CA certificate	Verifies a CM certificate in order to verify CM validity.	A manufacturer CA certificate is signed using the private key certificate of a root CA certificate.	The MA5600T/MA5603T/ MA5608T supports manufacturer CA certificate importing. When a CM goes online and registers with the MA5600T/MA5603T/ MA5608T, the MA5600T/MA5603T/ MA5608T uses the manufacturer CA certificate to verify the CM certificate and CM validity. This function protects the system from unauthorized CMs that may go online by impersonating authorized CMs.
CM certificate	A CM certificate is signed by a manufacturer for the CM. It is bound to the MAC address of a CM to uniquely prove the validity of the CM.	A CM certificate is signed using the private key of a manufacturer CA certificate.	The CM certificate and its private key file are installed on a CM when the CM is delivered. The CM does not provide an interface for reading data from or writing data to the private key file. This ensures the confidentiality of the private key file.

Verification Process

The certificate verification process meets the requirements of basic path validation defined in RFC 3280. The certificate verification process is as follows:

1. A CM reports its CM certificate and manufacturer CA certificate to the MA5600T/MA5603T/MA5608T.
2. The MA5600T/MA5603T/MA5608T uses the manufacturer CA certificate to verify the signature of the CM certificate. The MA5600T/MA5603T/MA5608T preferentially uses the manufacturer CA certificate imported to the system. If no manufacturer CA certificate has been imported to the system, the MA5600T/MA5603T/MA5608T uses the manufacturer CA certificate reported by the CM.
3. The MA5600T/MA5603T/MA5608T verifies the manufacturer CA certificate. The manufacturer CA certificate is signed using the root CA certificate. Therefore, the MA5600T/MA5603T/MA5608T must use the root CA certificate to verify the manufacturer CA certificate.
4. The MA5600T/MA5603T/MA5608T verifies the root CA certificate. The root CA certificate is signed by itself. Therefore, the root CA certificate can verify itself.

33.12.2 Configuring a Validity Check for a CM

The MA5600T/MA5603T/MA5608T supports the validity check for a CM feature. This feature prevents theft of network resources by an unauthorized CM.

Context

X.509 Digital Certificate

A CM must carry an X.509 digital certificate (a CM certificate), which contains the valid MAC address of the CM and uses the manufacturer CA certificate for the digital signature to verify the validity of the CM. A root CA certificate verifies the validity of a manufacturer CA certificate, enhancing the security of the MA5600T/MA5603T/MA5608T.

When a CM attempts to access a network, the MA5600T/MA5603T/MA5608T performs X.509 authentication for the digital certificate reported by the CM. This operation is performed to check whether the MAC address in the digital certificate is the same as that of the CM. If the MAC addresses are the same, the MA5600T/MA5603T/MA5608T considers the CM to be authorized and allows it to access the network. Otherwise, the MA5600T/MA5603T/MA5608T does not allow the CM to access the network.

Procedure

Run the **load certificate** command to load a root CA certificate or manufacturer CA certificate.

- root-ca: indicates the root CA certificate. This certificate is used to issue a manufacturer CA certificate. After it is imported, the system can check whether the manufacturer CA certificate is valid. Before a manufacturer CA certificate is authenticated, load the root CA certificate.

  The root CA certificate must be imported.
- maf-ca: indicates the manufacturer CA certificate. This certificate is used to issue a CM device certificate. After it is imported, the system can check whether the CM device certificate is valid. Before a CM certificate is authenticated, load the manufacturer CA certificate.

  The MA5600T/MA5603T/MA5608T preferentially uses the manufacturer CA certificate imported to the system. If no manufacturer CA certificate has been imported to the system, the MA5600T/MA5603T/MA5608T uses the manufacturer CA certificate reported by the CM.

----End

Example

The following is an example of the configurations used to load root and manufacturer CA certificates to authenticate the validity of a CM:

```
huawei(config)#load certificate root-ca tftp 10.10.10.10 CableLabs_CVC_Root_CA.cer
huawei(config)#load certificate mfg-ca tftp 10.10.10.10 CableLabs_CVC_CA.cer
```

33.13 Validity Check for a CM Configuration File

The MA5600T/MA5603T/MA5608T supports the validity check for a cable modem (CM) configuration file feature. This feature prevents theft of network resources by a CM that accesses the network using an unauthorized configuration file.

Introduction

After a CM is automatically discovered by the MA5600T/MA5603T/MA5608T, the CM obtains its configuration file from a Trivial File Transfer Protocol (TFTP) server. For details, see 33.3 CM Management. The MA5600T/MA5603T/MA5608T then allocates service flow resources to the CM based on the service flow, quality of service (QoS), and security configurations defined in the CM configuration file. After service flows are created between the MA5600T/MA5603T/MA5608T and the CM, user services can be forwarded between them.

CM configuration files can be compromised by two main methods:

- Method 1: An unauthorized user tampers with a CM configuration file to obtain service resources that they are not authorized to access.
- Method 2: An unauthorized user forges the CM configuration file of an authorized user to access a network and use network resources.

To protect against method 1, the MA5600T/MA5603T/MA5608T uses the validity check for a CM configuration file feature to verify the validity of the cable modem termination system (CMTS) message integrity check (MIC) in a CM registration message. This allows the MA5600T/MA5603T/MA5608T to confirm whether the configuration file obtained by the CM is from an authorized TFTP server.

To protect against method 2, the TFTP proxy is used to check whether the name of the CM configuration file is the same as that of the CM configuration file downloaded from the TFTP server to check whether the configuration file obtained by the CM is authorized by the TFTP server.

33.13.1 Principles

MIC

According to data over cable service interface specification (DOCSIS), the MA5600T/MA5603T/MA5608T needs to perform a message integrity check (MIC) on the registration request initiated by a CM.

The MA5600T/MA5603T/MA5608T performs an MIC check using either of the following MIC types: CM MIC or CMTS MIC.

- The CM MIC is used to check the completeness of a configuration file, that is, whether any configuration file data is lost during file transmission. It does not contain a private key.
- The CMTS MIC is used to check whether the TFTP server obtained at the CM registration stage is authorized. The CMTS MIC requires an encryption cipher key (CK). The MA5600T/MA5603T/MA5608T supports configuration of the encryption CK for the CMTS MIC.

The MA5600T/MA5603T/MA5608T obtains the CMTS MIC value using the HAMC-MD5 algorithm to calculate the fields that need to be reported in a registration message and that are defined in the CM configuration file. The MA5600T/MA5603T/MA5608T then checks whether the CMTS MIC value obtained from the registration message is the same as the calculated CMTS MIC value. If the values are different, the MIC check fails and the CM fails to register with the MA5600T/MA5603T/MA5608T.

33.13.2 Configuring a Validity Check for a CM Configuration File

The MA5600T/MA5603T/MA5608T supports the validity check for a CM configuration file feature. This feature prevents theft of network resources by an unauthorized CM.

Context

MIC

An MIC is used to check whether the configuration file obtained by the CM is from an authorized TFTP server. This operation prevents an unauthorized user from tampering with the CM configuration file to obtain service resources that they are not authorized to access.

The MA5600T/MA5603T/MA5608T uses the HAMC-MD5 algorithm to calculate the fields that need to be reported in a registration message to obtain the CMTS MIC value. The MA5600T/MA5603T/MA5608T then checks whether the CMTS MIC value obtained from the registration message is the same as the calculated CMTS MIC value.

- If the values are different, the MIC check fails and the CM fails to register with the MA5600T/MA5603T/MA5608T.
- If the values are the same, the MIC check is successful and the CM configuration file is authorized.

Procedure

- Configure the MIC check.
  a. Run the **cable shared-secret** { **simple** | **cipher** } *authentication-key* command to set the encryption key of an MA5600T/MA5603T/MA5608T MIC check.
     - **simple**: indicates the plaintext key. After this parameter is used, the password is encrypted and saved in the configuration file. The encrypted password cannot be viewed, ensuring high security.
     - **cipher**: indicates the ciphertext key. A ciphertext key is not obtained through this parameter but through an HAMC-MD5 algorithm.

----End

Example

The MIC check needs to be configured to check the validity of a CM configuration file to enhance network security. In the example provided in this document, the key for the MIC check is **huawei123**.

```
huawei(config-if-cable-0/1/0)#cable shared-secret simple huawei123
```

33.14 Built-in Optical Transceiver

Introduction

Background

In traditional HFC networks, VoD is implemented as follows: Upstream signals are transmitted using the IP technology over data channels; downstream inband RF signals are transmitted at a fixed frequency specified by the MSO from the head end video system to STBs. The downstream inband RF signals also carry EPG and conditional access (CA) data.

Out of band (OOB) is a technology for exchanging data between the VoD service and the digital video service. In OOB technology, a bidirectional data channel independent of traditional DOCSIS channels is set up to transmit and receive digital TV data, including but not limited to CA, service information (SI), EPG, electronic article surveillance (EAS), and VoD interaction signaling in both downstream and upstream directions. This technology requires dedicated OOB STBs and is used in some countries.

Huawei has provided the solution of external optical transceiver+MA5633 for OOB applications. However, the external optical transceiver, an independent device, not only complicates networking but also increases installation and maintenance costs. To resolve the issues caused by external transceivers, Huawei has developed the MA5633 equipped with a built-in optical transceiver.

Benefits

The MA5633 equipped with a built-in optical transceiver brings the following benefits to carriers:

- **Reduces optical fiber costs**: The built-in optical transceiver transmits and receives broadband and video data over one optical fiber, without changing existing feeder fiber deployment.
- **Reduces installation costs**: The built-in optical transceiver is integrated inside the MA5633, removing the need of cable connections and commissioning and that of renting installation space required by the external transceiver. This design simplifies installation and reduces cabinet space on an FN.
- **Reduces failure rates**: The built-in optical transceiver reduces the risk of signal quality deterioration caused by insecure cable connections. In addition, the D-CCAP supports remote management and maintenance for the built-in optical transceiver and provides data backhaul for traditional cable analyzers.

33.14.2 Principles

A built-in optical transceiver consists of an optical receiver and an optical transmitter.

- The optical receiver receives optical signals transmitted over optical fibers in the CATV transmission system, converts optical signals to electrical signals, and outputs RF signals.
- The optical transmitter provides OOB signal backhaul for the VoD service. It uses the optical transmission backhaul module built in the MA5633 to send OOB data in upstream frequency bands to the head end video system for demodulation.

33.14.3 Usage Scenarios

Networking Introduction

Figure 33-58 shows the networking of an MA5633 equipped with a built-in optical receiver. Figure 33-59 shows the networking of an MA5633 equipped with a built-in optical transceiver for OOB applications.

Figure 33-58 Networking of a D-CAAP equipped with a built-in optical receiver

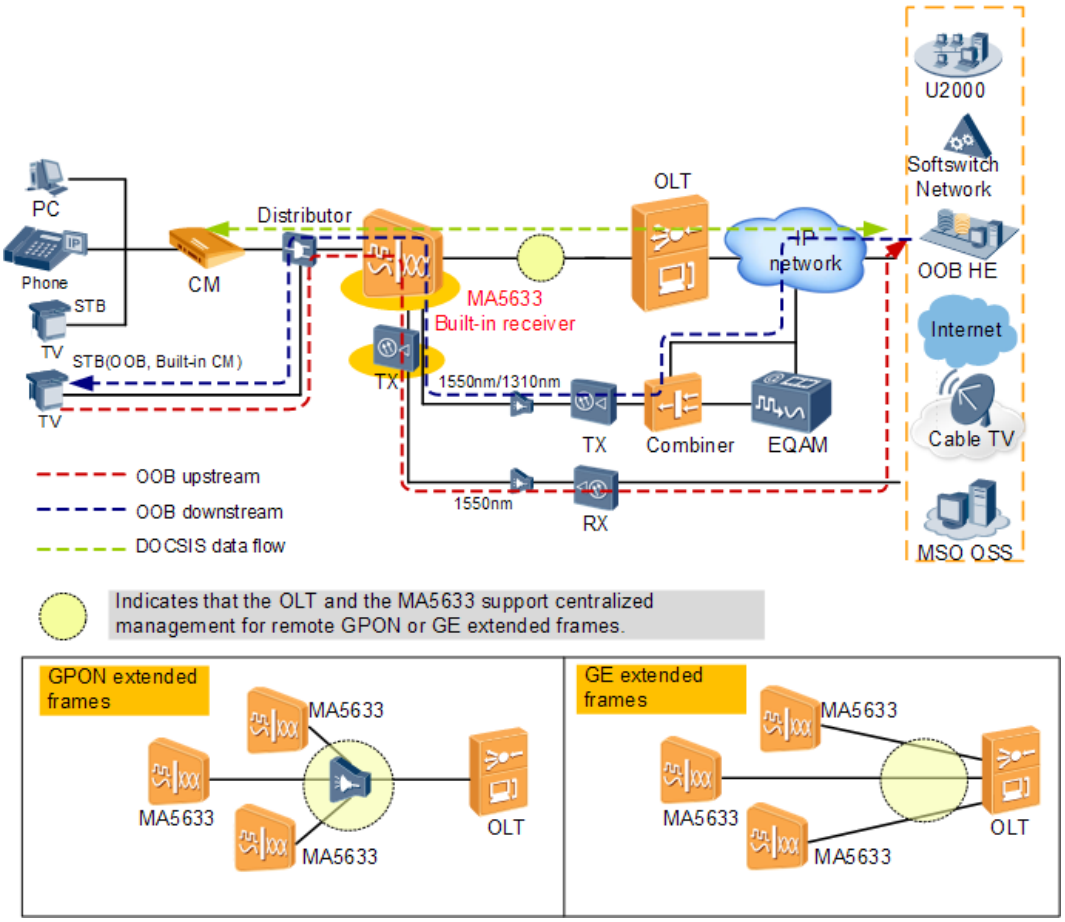
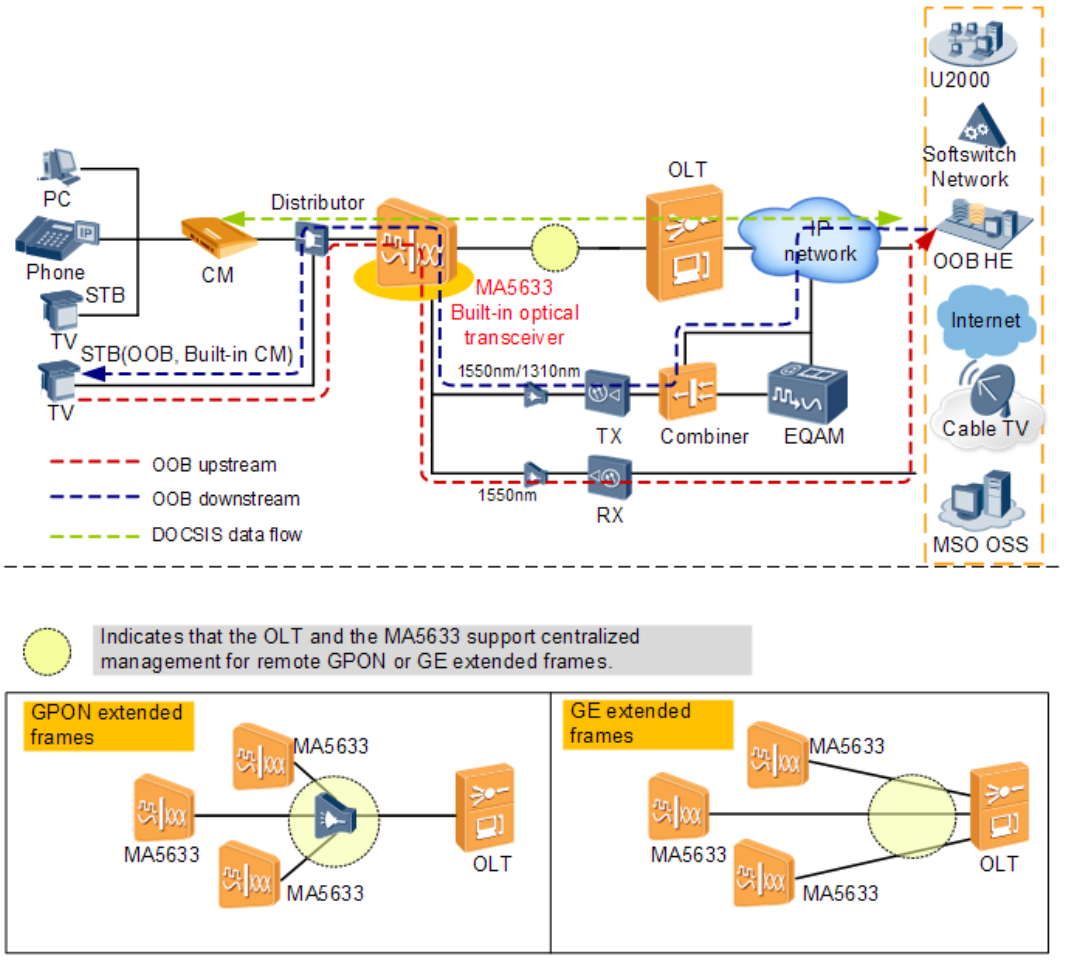


Figure 33-59 Networking of a D-CAAP equipped with a built-in optical transceiver for OOB applications



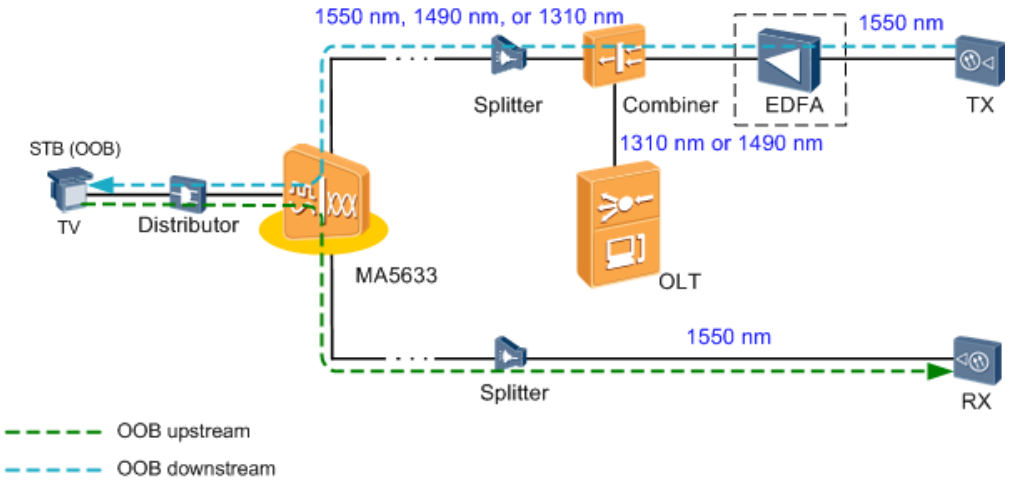
A built-in optical transceiver simplifies a traditional OOB network as follows:

- In the downstream direction, digital TV and VoD signals are transmitted with CATV signals to the HFC network in a dedicated OOB frequency band. The built-in optical receiver module of the MA5633 transparently transmits these signals to OOB STBs at user homes.
- In the upstream direction, OOB STBs send OOB signals through the optical transmission backhaul module built in the MA5633 to the head end OOB server in the upstream frequency band for demodulation.

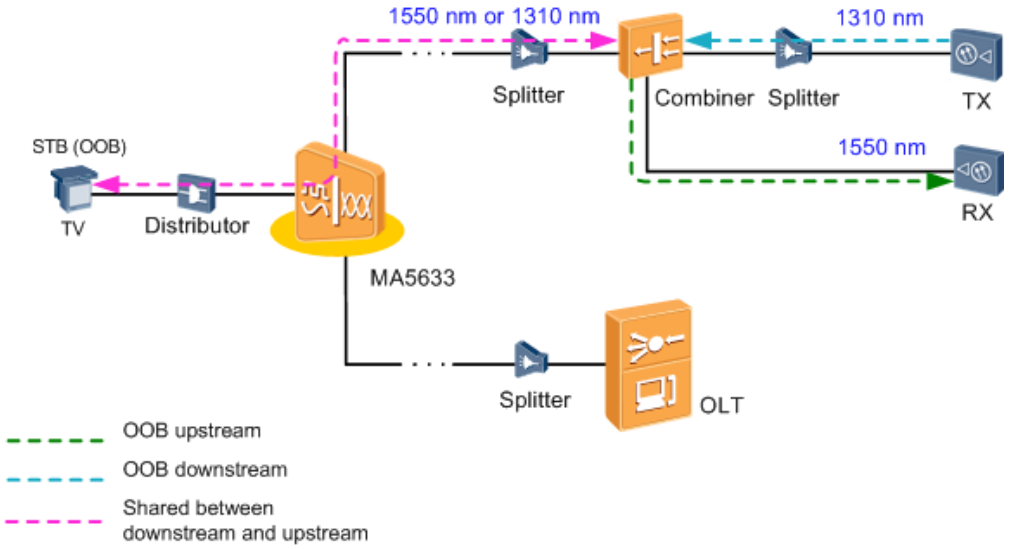
Typical Usage Scenarios

Built-in optical transceiver networking varies depending on the transmission of downstream and upstream signals and optical operating wavelengths. The following section describes three typical usage scenarios of built-in optical transceivers.

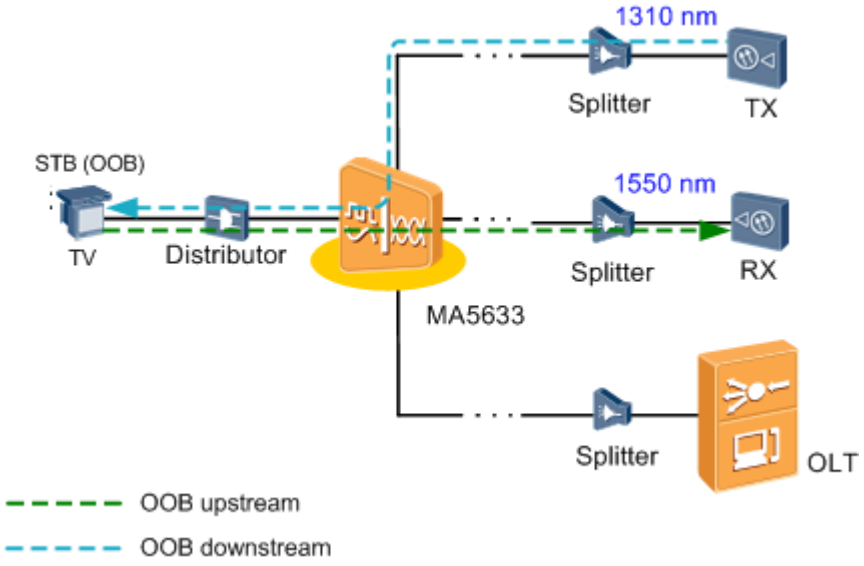
- Typical usage scenario 1: Downstream and upstream OOB signals are transmitted over separate optical fibers; both downstream and upstream center wavelengths are 1550 nm; OOB signals and data signals are combined in the downstream direction using a combiner.



- Typical usage scenario 2: Downstream and upstream OOB signals are transmitted over one optical fiber after being combined using a combiner; downstream and upstream center wavelengths are 1310 nm and 1550 nm, respectively; OOB signals and data signals are transmitted over separate optical fibers.



- Typical usage scenario 3: Downstream and upstream OOB signals are transmitted over separate optical fibers; downstream and upstream center wavelengths are 1310 nm and 1550 nm, respectively; OOB signals and data signals are transmitted over separate optical fibers.



33.14.4 Maintenance and Diagnosis

Indicators

The MA5600T/MA5603T/MA5608T provides indicators for the built-in optical transceiver so that users can obtain the running status of the built-in optical transceiver. Table 33-23 lists the indicators for the built-in optical transceiver and their meanings.

Table 33-23 Indicators for the built-in optical transceiver and their meanings

Silkscreen	Status	
POWER	Steady green	The power supply to the built-in optical transceiver is normal.
	Green off	The power supply to the built-in optical transceiver is abnormal.
RX	Steady green	The optical receiver is working properly. Specifically, the input optical power is within the allowed range.
	Steady red	The optical receiver malfunctions. The possible causes are as follows: The input optical power is greater than the upper threshold or less than the lower threshold for the optical power supported by the optical receiver, or no optical power is input into the optical receiver.
TX	Steady green	The optical transmitter is working properly. Specifically, the output optical power is within the allowed range.
	Steady red	The optical transmitter malfunctions. The possible causes are as follows: The output optical power is greater than the upper threshold or less than the lower threshold for the optical power supported by

Silkscreen	Status	
		the optical transmitter, or the optical transmitter outputs no optical power.

Alarms

The alarms reported for a built-in optical transceiver can be used to locate a fault.

Table 33-24 List of alarms related to a built-in optical transceiver

Alarm ID	Alarm Name
0x1541121d	Abnormal RF power of the optical-receiver
0x1531a001	The communication of the optical-node with the device is abnormal
0x1541121c	Abnormal optical power of the optical-receiver
0x1541121f	Abnormal bias current of the optical-transmitter
0x1541121e	Abnormal optical power of the optical-transmitter

33.14.5 Standards and Protocols Compliance

Table 33-25 lists the standards and protocols that a built-in optical transceiver complies with.

Table 33-25 Standards and protocols that a built-in optical transceiver complies with

Standard and Protocol	Description
ANSI/SCTE-55-1 (formerly DVS 178)	Digital Broadband Delivery System: Out of band Transport Part 1 — Mode A
ANSI/SCTE-55-2 (formerly DVS 167)	Digital Broadband Delivery System: Out of band Transport Part 2 — Mode B

33.15 Spectrum Management

The spectrum management feature enables the MA5600T/MA5603T/MA5608T to minimize noise interference on signals transmitted over upstream channels on hybrid fiber coaxial (HFC) networks. This ensures the quality of service (QoS) of cable users' data and voice services.

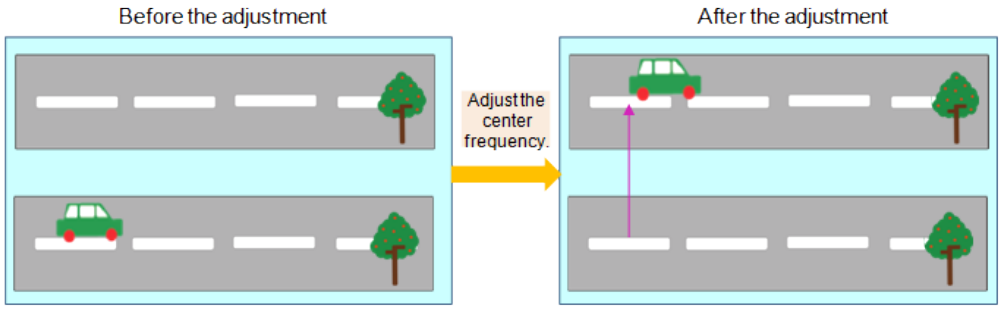
What Are Spectrum Management Policies

Upstream channels of distributed converged cable access platform (D-CCAP) are prone to be interfered by upstream noises, degrading the transmission quality of upstream channels and accordingly degrading the quality of cable users' services. Spectrum management policies are a series of adjustment policies preventing noises from interfering with upstream channels.

Specifically, when the increase of channel interference degrades signal transmission, the spectrum management policies configured on the D-CCAP allow the D-CCAP to adjust the center frequency, frequency bandwidth, or modulation profile to prevent noise interference, thereby improving channel transmission quality. Spectrum management policies involve the following parameters:

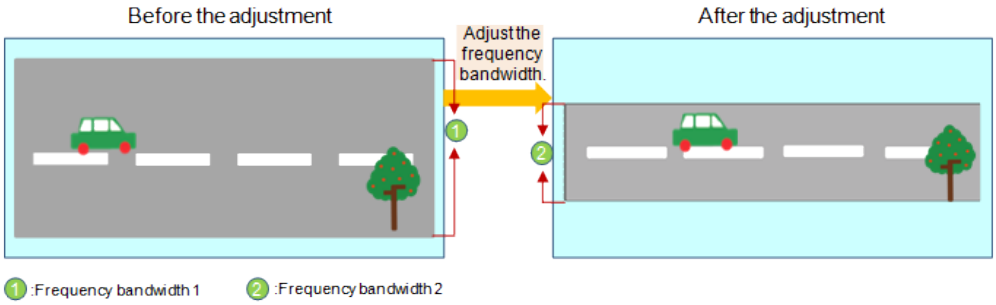
- Center frequency: center of a channel spectrum, which is similar to a road center line. The center frequency and frequency bandwidth of a channel determine the channel frequency range. Figure 33-60 shows the adjustment of a center frequency.

  Figure 33-60 Adjusting a center frequency

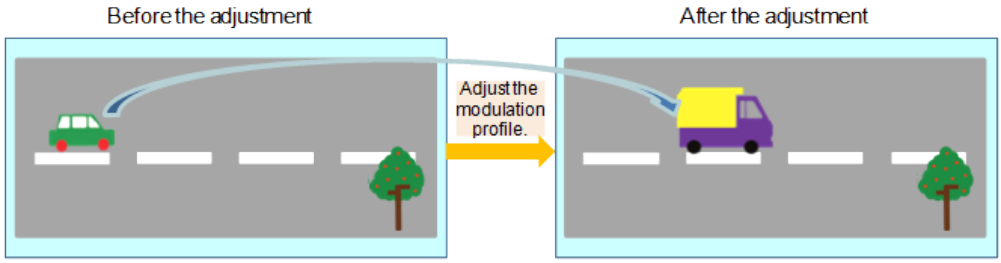
- Frequency bandwidth: specifies the spectrum width of a channel, which is similar to a road width. Figure 33-61 shows the adjustment of a frequency bandwidth.

  Figure 33-61 Adjusting a frequency bandwidth

- Modulation profile: specifies the parameters required for processing signals in upstream channels. Modulation profiles vary depending on signal-to-noise ratios (SNRs). Therefore, the modulation profile must be adjusted for a channel when the SNR of this channel is changed out of the range supported by this modulation profile. This ensures signal transmission with expected quality. The rules of adjusting a modulation profile are the same as those of determining a transport mode based on road conditions for smooth destination arrival. Figure 33-62 shows the adjustment of a modulation profile.

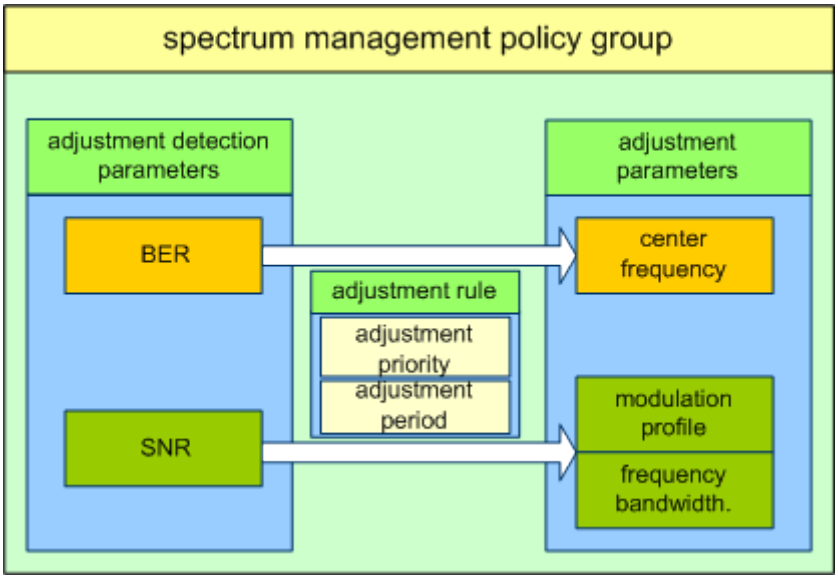
Figure 33-62 Adjusting a modulation profile



33.15.2 Basic Concepts in the Spectrum Management Policy

The Distributed-Converged Cable Access Platform (D-CCAP) supports the binding of a spectrum management policy group to an upstream channel. The spectrum management policy group defines adjusted parameters (center frequency, frequency bandwidth, and modulation profile), detected parameters, and an adjustment rule. As shown in Figure 33-63, when detected parameters are higher than the thresholds (that is, the channel transmission quality deteriorates), the D-CCAP adjusts parameters based on the adjustment rule. By doing so, the D-CCAP monitors and manages the frequency spectrum of the channel, preventing noises from affecting services.

Figure 33-63 Mapping between detected parameters and adjusted parameters



Detected Parameters Adjustment

The D-CCAP determines channel quality by checking the bit error rate (BER) and signal-to-noise ratio (SNR) of a channel.

- BER: A higher BER represents a poorer transmission quality of a channel. After detecting that the BER is higher than the threshold, the D-CCAP adjusts only the center frequency.

- SNR: A lower SNR represents a poorer transmission quality of a channel. After detecting that the SNR is lower than the threshold, the D-CCAP adjusts only the frequency bandwidth and modulation profile.

Adjusted Parameters

Table 33-26 lists adjusted parameters.

Table 33-26 Relationship between detected parameters and adjusted parameters

Detected Parameter	Adjusted Parameter	Adjusted Parameter Description	Adjustment Rule
BER	Center frequency	Each spectrum management policy group supports a maximum of four backup center frequencies. The four backup center frequencies must fall within the frequency range defined in the corresponding European or North American standard. Each backup center frequency must be unique and cannot be the same as the original center frequency of the upstream channel.	When the BER is higher than the threshold, the D-CCAP adjusts only the center frequency. The D-CCAP uses the backup center frequency configured earlier as the new center frequency.
SNR SNR thresholds are as follows: • Level-1 SNR threshold • Level-2 SNR threshold **NOTE** The level-1 SNR threshold must be higher than the level-2 SNR threshold	Frequency bandwidth	Each spectrum management policy group supports only one backup frequency bandwidth. The backup frequency bandwidth must be lower than the original frequency bandwidth of the upstream channel. **NOTE** The frequency bandwidth adjustment reduces channel rates.	The D-CCAP adjusts the frequency bandwidth only after detecting that the SNR is lower than the level-2 SNR threshold.
	Modulation profile	Each spectrum management policy group supports two backup modulation profiles: level-1 backup modulation profile and level-2 backup modulation profile. A backup modulation profile can only be one of the modulation profiles pre-configured on the D-CCAP.	• If the detected SNR is higher than the level-2 SNR threshold and lower than the level-1 SNR threshold, the D-CCAP sets the level-1 backup modulation profile as the modulation profile. • If the detected SNR is lower than the level-2 SNR threshold, the D-CCAP adjusts the modulation

Detected Parameter	Adjusted Parameter	Adjusted Parameter Description	Adjustment Rule
.			profile to the level-2 backup modulation profile.

Adjustment Rule

Adjustment priority

If the detected BER of an upstream channel is higher than a threshold or the SNR is lower than a threshold, the D-CCAP adjusts parameter settings according to the adjustment priority configured in the frequency spectrum management policy group. The D-CCAP supports the following three types of adjustment priorities:

- Center frequency > modulation profile > frequency bandwidth
- Center frequency > frequency bandwidth > modulation profile
- Modulation profile > Center frequency > frequency bandwidth

The D-CCAP automatically records parameter adjustment logs. Both the previously adjusted parameter and the adjustment priority determine the parameter to be adjusted. The following provides an example to describe the adjustment path:

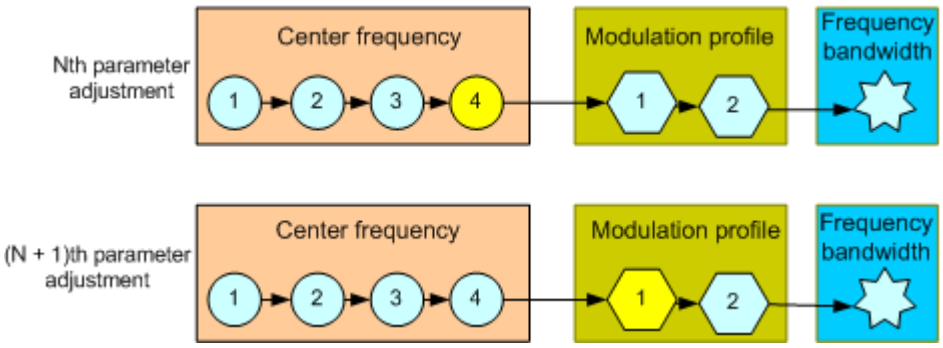
- Adjustment priority is center frequency > modulation profile > frequency bandwidth.
- Four backup center frequencies are configured, frequencies 1, 2, 3, and 4.
- Two backup modulation profiles are configured, modulation profiles 1 and 2.
- One frequency bandwidth is configured.

In this situation, the adjustment path is frequency 1 > frequency 2 > frequency 3 > frequency 4 > modulation profile 1 > modulation profile 2 > frequency bandwidth. As shown in Figure 33-64, the previously adjusted parameter is frequency 4. Then, the parameter to be adjusted is modulation profile 1.

NOTE

For a modulation profile, if the detected SNR is lower than the level-2 SNR threshold, the D-CCAP adjusts modulation profile 2, even if modulation profile 1 should be adjusted according to the adjustment rule.

Figure 33-64 Adjustment path



Adjustment period

Except the center frequency, the adjustment for other parameters causes CMs to go offline and then online. Frequent parameter adjustment leads to frequent going offline for CMs, interrupting user services. The spectrum management policy group defines the parameter adjustment protection period and parameter adjustment cycling period, preventing frequent adjustments.

- Parameter adjustment protection period: Parameter adjustment can be performed only once within the period. The period is configurable and 30 minutes by default.
- Parameter adjustment cycling period: All parameters can be adjusted only once within the period. The default period of 24 hours is used.

33.15.3 Spectrum Management Principles

After a spectrum management policy group is bound to an upstream channel, the distributed converged cable access platform (D-CCAP) adjusts the center frequency, frequency bandwidth, or modulation profile of the upstream channel based on the rule defined in the spectrum management policy group. The adjustment process varies depending on scenarios. For details, see Table 33-27.

Table 33-27 Scenarios and corresponding adjustment processes

Scenario	Adjustment Process
The D-CCAP detects only the bit error rate (BER), that is, the D-CCAP adjusts only the center frequency.	For details, see BER Detection and Center Frequency Adjustment Process.
The D-CCAP detects only the signal-to-noise ratio (SNR) and adjusts only the modulation profile.	For details, see SNR Detection and Modulation Profile Adjustment Process.
The D-CCAP detects only the SNR and adjusts only the frequency bandwidth. **NOTE** Adjusting the frequency bandwidth is not recommended because it decreases channel rates.	For details, see SNR Detection and Frequency Bandwidth Adjustment Process.
The D-CCAP adjusts all parameters. **NOTE**	For details, see Example for Adjusting All Parameters.

Scenario	Adjustment Process
For scenarios in which only some parameters are adjusted, configure only required parameters. If the priority sequence required by the actual situation differs from that in the given configuration example, adjust the priority sequence.	
The SNR increases after the modulation profile is adjusted.	For details, see Modulation Profile Switchback.
Channel quality deteriorates after all parameters are adjusted.	For details, see Parameter Adjustment Restoration.

NOTE

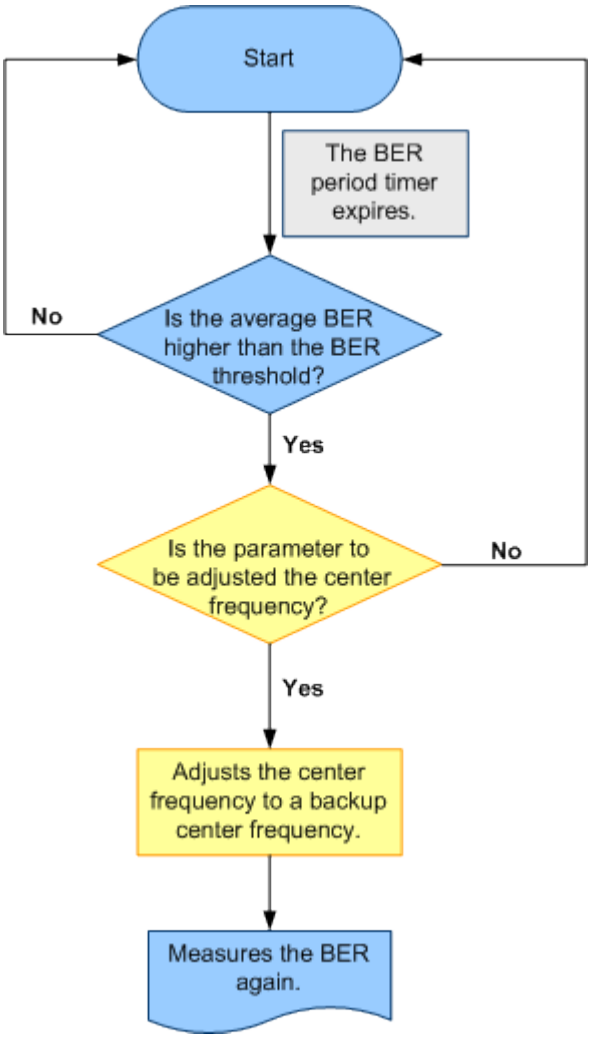
Except the parameter adjustment restoration, other processes can be performed only once within the adjustment protection period, and all parameters can be adjusted only once within 24 hours.

BER Detection and Center Frequency Adjustment Process

As shown in Figure 33-65, when the BER period timer expires, the D-CCAP determines whether the average BER within the detection period is higher than the preset threshold. The BER period timer duration is 15 minutes.

- If the average BER is higher than the threshold and the parameter to be adjusted is a center frequency, the D-CCAP adjusts the center frequency. If the parameter to be adjusted is not a center frequency, the D-CCAP starts another BER detection.
- If the average BER is lower than the threshold, the D-CCAP starts another BER detection.

Figure 33-65 BER detection and center frequency adjustment process



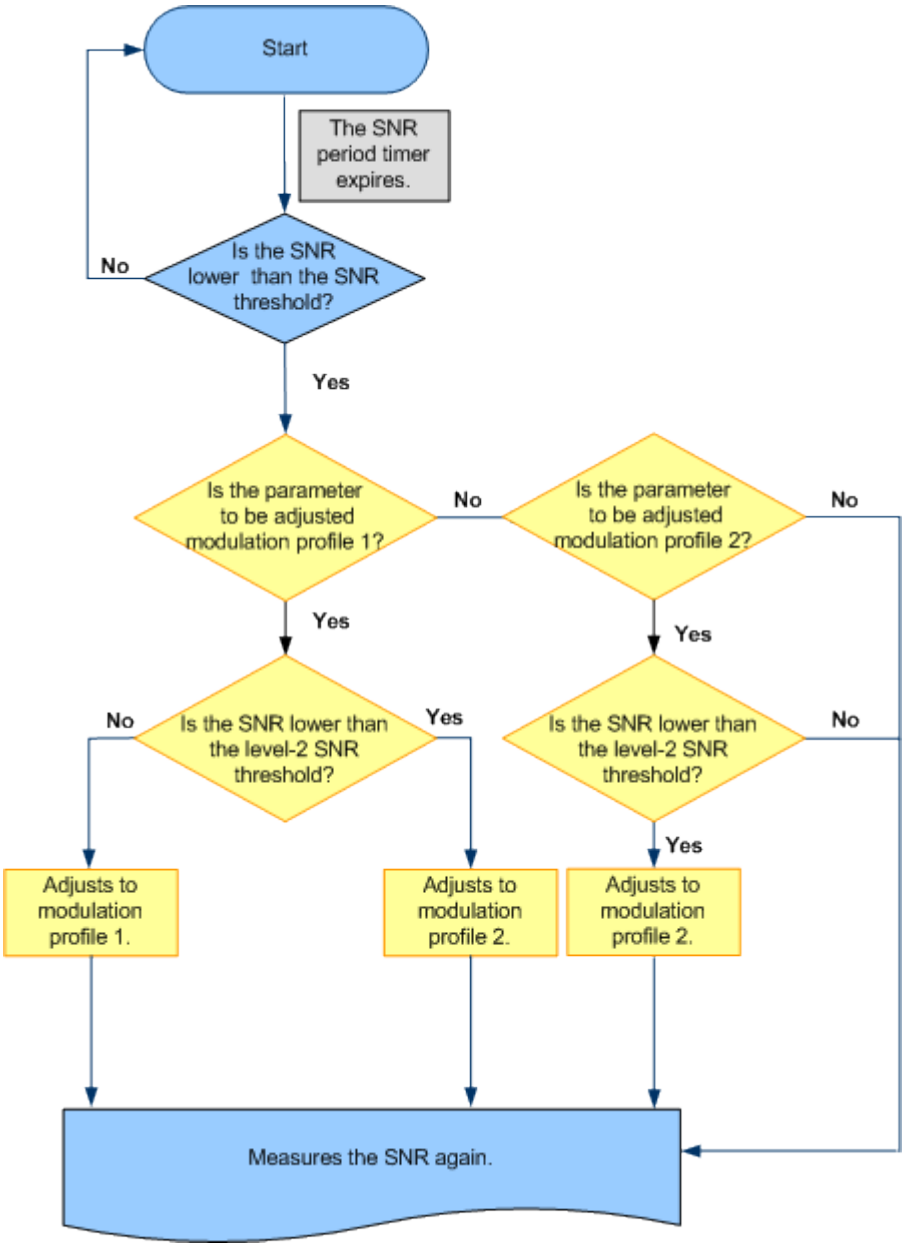
SNR Detection and Modulation Profile Adjustment Process

As shown in Figure 33-66, after the SNR period timer expires (timer duration is configurable), the D-CCAP performs the following checks:

1. Checks whether the detected SNR is lower than the level-1 or level-2 SNR threshold.
   - If the detected SNR is lower than either level-1 or level-2 SNR threshold, the D-CCAP determines the parameter to be adjusted. If the parameter to be adjusted is modulation profile 1, the D-CCAP performs the operation in step 2. If the parameter to be adjusted is modulation profile 2, the D-CCAP performs the operation in step 3. If other parameters are to be adjusted, the D-CCAP starts another SNR detection.
   - If the average SNR is higher than the threshold, the D-CCAP starts another SNR detection.
2. Checks whether the detected SNR is lower than the level-2 SNR threshold. If the detected SNR is lower than the level-2 SNR threshold, the D-CCAP adjusts modulation profile 2. If the detected SNR is higher than the level-2 SNR threshold but lower than the level-1 threshold, the D-CCAP adjusts modulation profile 1.

3. Checks whether the detected SNR is lower than the level-2 SNR threshold. If the detected SNR is lower than the level-2 SNR threshold, the D-CCAP adjusts modulation profile 2. Otherwise, the D-CCAP starts another SNR detection.

Figure 33-66 SNR detection and modulation profile adjustment process



SNR Detection and Frequency Bandwidth Adjustment Process

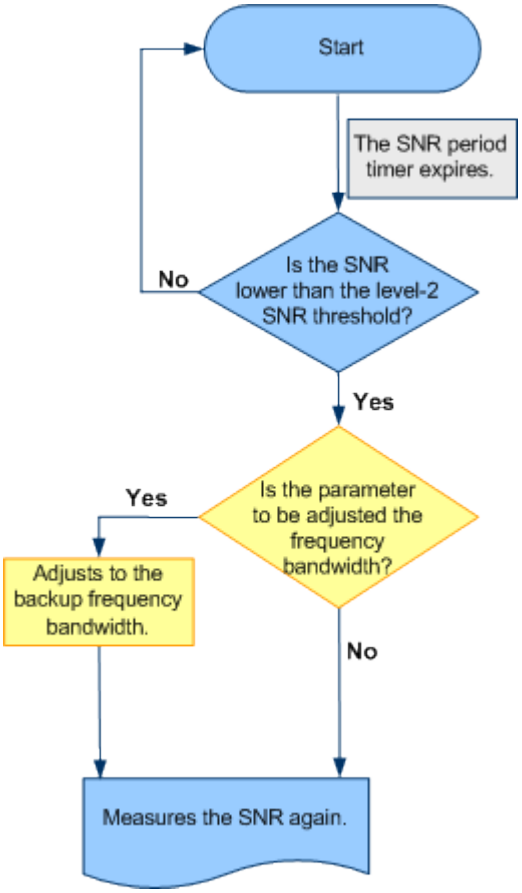
As shown in Figure 33-67, after the SNR period timer expires (timer duration is configurable), the D-CCAP checks whether the detected SNR is lower than the level-2 SNR threshold.

- If the detected SNR is lower than the level-2 SNR threshold, the D-CCAP determines the parameter to be adjusted. If the parameter to be adjusted is frequency bandwidth, the

D-CCAP adjusts the frequency bandwidth. If other parameters are to be adjusted, the D-CCAP starts another SNR detection.

- If the average BER is higher than the level-2 SNR threshold, the D-CCAP starts another SNR detection.

Figure 33-67 SNR detection and frequency bandwidth adjustment process



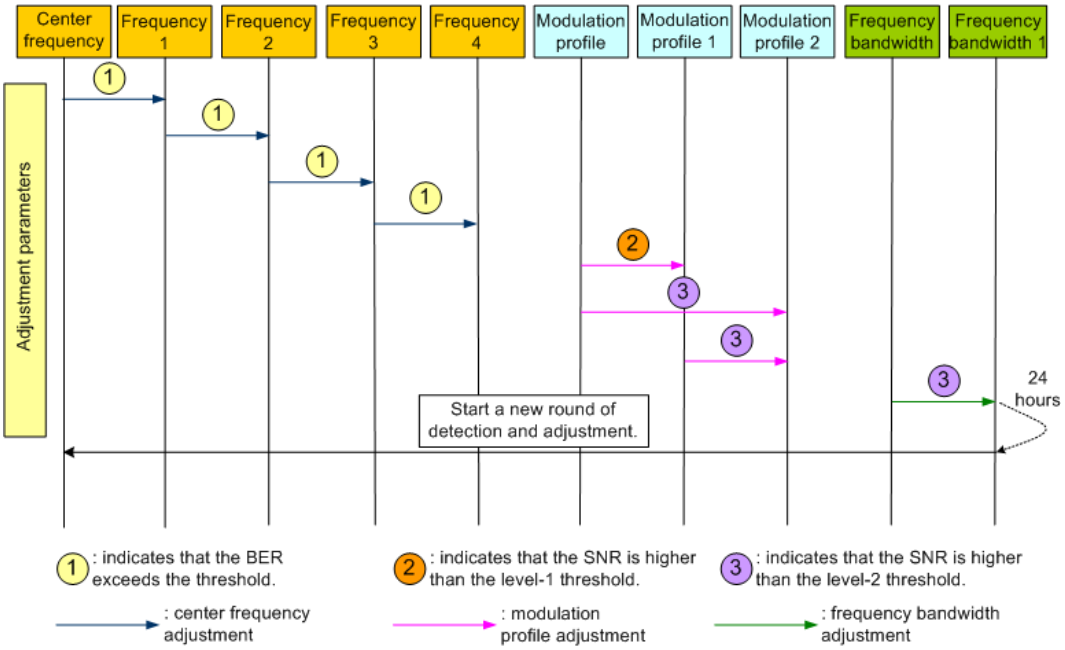
Example for Adjusting All Parameters

Parameter configurations are as follows:

- Adjustment priority: center frequency > modulation profile > frequency bandwidth
- Four backup center frequencies 1, 2, 3, and 4
- Two backup modulation profiles: level-1 modulation profile 1 and level-2 modulation profile 2
- Backup frequency bandwidth 1

As shown in Figure 33-68, the D-CCAP adjusts parameters based on the configured priority.

Figure 33-68 Adjustment process for all parameters

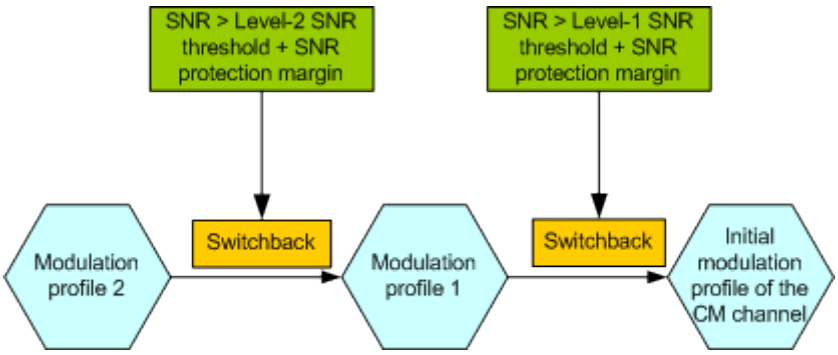


Modulation Profile Switchback

Modulation profile switchback is performed when the channel quality recovers and the SNR is higher than the specified SNR threshold. Frequent adjustments affect services on CMs. To avoid this issue, a modulation profile switchback is allowed only when the SNR is higher than the sum of an SNR threshold and the SNR protection margin.

A modulation profile can only be switched back to the previous module profile. In addition, a switchback does not change the parameter adjustment path. After a switchback, the parameter adjustment path continues from the position before the switchback and follows the original priority. Figure 33-69 shows a modulation profile switchback.

Figure 33-69 Modulation profile switchback



Parameter Adjustment Restoration

After adjusting all parameters in a frequency spectrum management policy group, restore the original settings if the upstream channel quality is poorer than that before the adjustment. To

do so, run the **undo cable upstream hop** command to unbind the spectrum management policy group from the upstream channel. Then, the center frequency, modulation profile, and frequency bandwidth of the upstream channel will be restored to their original settings. After the parameter adjustment restoration, all CMs connected to the upstream channel go offline and online again.

33.15.4 Configuring a Spectrum Management Policy Group

This section describes how to configure a spectrum management policy group. After the configuration, distributed converged cable access platform (D-CCAP) can adjust a center frequency, frequency bandwidth, or modulation profile for data transmission based on the spectrum management policy. This prevents upstream noise from affecting services.

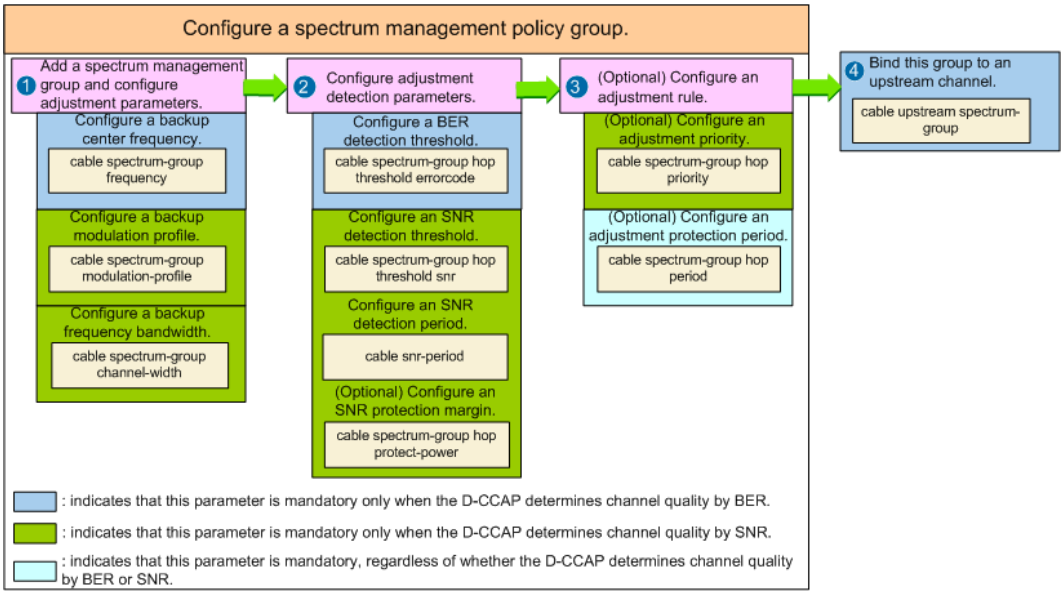
Prerequisites

RF parameters have been configured for the upstream channel, and the upstream channel has been activated.

Context

Figure 33-70 shows the process for configuring the spectrum management policy.

Figure 33-70 Process for configuring the spectrum management policy



Procedure

Add a spectrum management policy group and configure adjustment parameters.

New spectrum management policy groups are accordingly added when backup center frequencies, backup frequency bandwidths, or backup modulation profiles are added.

1. Run the **cable spectrum-group frequency** command to configure a center frequency in the spectrum management policy group.

Each spectrum management policy group supports a maximum of four backup center frequencies. If the BER exceeds the threshold, the D-CCAP adjusts the center frequency according to the frequency configuration sequence. The D-CCAP supports a maximum of four FH attempts every 24 hours.

1. Run the **cable spectrum-group modulation-profile** command to configure a backup modulation profile in the spectrum management policy group.

   Each spectrum management policy group supports two backup modulation profiles: level-1 backup modulation profile and level-2 backup modulation profile.

2. Run the **cable spectrum-group channel-width** command to configure a backup frequency bandwidth in the spectrum management policy group.

   Each spectrum management policy group supports only one backup frequency bandwidth. A backup frequency bandwidth and a center frequency of an upstream channel determine the channel frequency range. Ensure that the new frequency range does not overlap that of other channels.

Step 1 Configure adjustment detection parameters.

1. Run the **cable spectrum-group hop threshold** command to configure the BER and SNR detection thresholds in the spectrum management policy group.
   - If the detected BER is higher than the threshold, the D-CCAP adjusts the center frequency.
   - If the detected SNR is lower than the threshold, the D-CCAP adjusts the frequency bandwidth or modulation profile according to the adjustment rule.
2. Run the **cable snr-period** command to configure the SNR detection period in the spectrum management policy group.

   The SNR is detected only once within a detection period. The D-CCAP does not detect the SNR by default.
3. (Optional) Run the **cable spectrum-group hop protect-power** command to configure an SNR protection margin in the spectrum management policy group. The default value is 3 dB. If the default value cannot meet requirements, change it.

   If the sum of the detected SNR and the SNR protection margin is greater than the SNR threshold, the D-CCAP switches back modulation profiles level by level.

Step 2 (Optional) Configure an adjustment rule.

1. (Optional) Run the **cable spectrum-group hop priority** command to configure an adjustment priority in the spectrum management policy group.

   By default, the adjustment priority in the spectrum management policy group is center frequency > modulation profile > frequency bandwidth. Configure the adjustment priority if the default priority cannot meet requirements.
2. (Optional) Run the **cable spectrum-group hop period** command to configure an adjustment protection period in the spectrum management policy group. The frequency spectrum can be changed only once within a period.

   The default adjustment protection period is 1800s. Change the adjustment protection period if the default value cannot meet requirements.

Step 3 Bind the spectrum management policy group to an upstream channel.

Run the **interface cable** command to enter cable mode, and then run the **cable upstream spectrum-group** command to configure the spectrum management policy group bound to an upstream channel. Each channel supports only one spectrum management policy group. When binding a spectrum management policy group to an upstream channel, ensure that the frequency spectrum used by the channel does not overlap that used by other channels.

Step 4 (Optional) Run the **display cable spectrum-group** command to check whether the data configured in the spectrum management policy group complies with the planned data.

Step 5 (Optional) Run the **display cable upstreamhop trace** command to query the parameter adjustment path of the upstream channel.

In maintenance scenarios, you can query D-CCAP parameter adjustment logs to obtain information about past parameter setting changes. This facilitates fault locating. The D-CCAP records a maximum of 30 parameter adjustment logs.

----End

Example

The following is an example of the configurations used to bind spectrum management policy group 1 to upstream channel 1:

- Center frequencies: 7 MHz, 8.6 MHz, 10.2 MHz, and 11.8 MHz
- BER detection threshold: 20%
- FH protection period: 9600s
- Level-1 SNR threshold: 28 dB; level-2 SNR threshold: 18 dB
- SNR detection period: 1000 ms
- Frequency bandwidth: 1600 kHz
- Level-1 backup modulation profile: pre-configured modulation profile 1; level-2 backup modulation profile: pre-configured modulation profile 2

```
huawei(config)#cable spectrum-group 1 frequency 7 8.6 10.2 11.8
huawei(config)#cable spectrum-group 1 hop threshold errorcode 20 snr 28 18
huawei(config)#cable spectrum-group 2 hop period 9600
huawei(config)#cable snr-period 1000
huawei(config)#cable spectrum-group 1 modulation-profile secondary-modulation-profi
le1 2 secondary-modulation-profile2 2
huawei(config)#cable spectrum-group 1 channel-width 1600
huawei(config)#display cable spectrum-group 1
  ---------------------------------------------------------------------------
  Spectrum Management Policy Group ID : 1
  Central Frequency(MHz)            : 7.00,8.60,10.20,11.80
  FH Detection Threshold(%)         : BER (Average BER threshold = 20)
  FH Protection Period(s)           : 9600
  BER Detection Threshold(%)        : 5
  Secondary Modulation Profile1     : 2
  Secondary Modulation Profile2     : 4
  Channel Width(KHz)                : 1600
  First SNR Threshold(dB)           : 28
  Second SNR Threshold(dB)          : 18
  Protect Power(dB)                 : 3
  Priority                          : Frequency > Modu-Profile > Chan-Width
  ---------------------------------------------------------------------------
huawei(config)#interface cable 1/1/0
huawei(config-if-cable-1/1/0)#cable upstream 1 spectrum-group 1
```

33.16 IPDR

The Internet Protocol Detail Record (IPDR) feature enables the MA5600T/MA5603T/MA5608T to collect data, encode the collected data in an IPDR-dedicated external data representation (XDR) format, and send the encoded data to an IPDR server. The collected data includes accounting information, running statuses of the cable modem (CM) served by the MA5600T/MA5603T/MA5608T, frequency spectrum information, debugging information, and statistics for cable modem termination system (CMTS) bandwidths and service flows. The IPDR feature complies with data over cable service interface specification (DOCSIS) 2.0 and 3.0.

What Is IPDR

Background

The PON+distributed converged cable access platform (D-CCAP) solution is widely used by multiservice operators (MSOs) and is serving an increasing number of users. When used in a PON+D-CCAP solution, the MA5600T/MA5603T/MA5608T can collect only limited data about user accounting information and network running status.

The IPDR feature resolves these issues by providing a more effective data collection and reporting mechanism for carriers through an operations support system (OSS).

Typical Application

The data collected using the IPDR feature can be used to:

- Manage and monitor service data based on service type.
- Configure the CMTS head end capacity based on service traffic.
- Control and manage traffic based on the traffic volumes of different service types.
- Analyze the MSO networks and user behavior based on service flows and user application information.

NOTE

Specific usage of the data collected using the IPDR feature is subject to the carrier's planning and therefore is not described in this document.

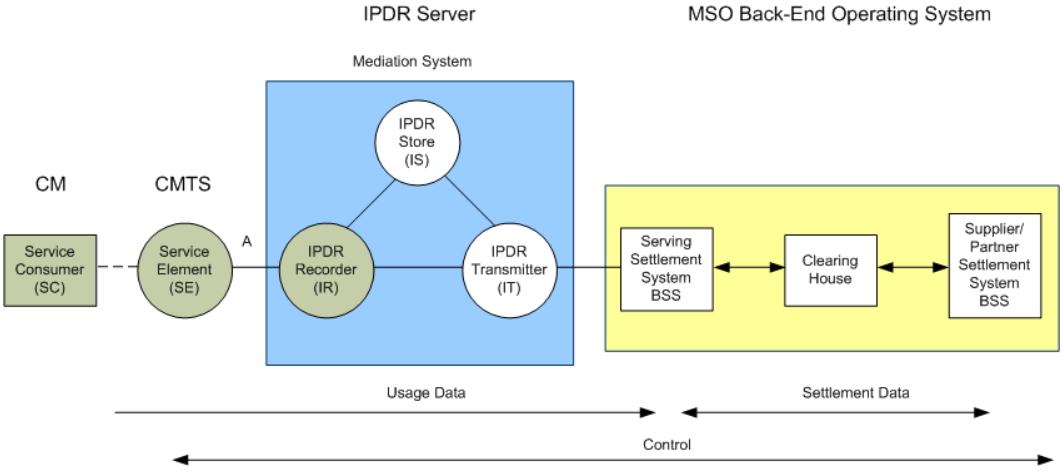
33.16.2 Basic IPDR Concepts

This section describes basic IPDR concepts.

IPDR Model

According to DOCSIS, a basic IPDR network model consists of a service consumer, a service element, a mediation system, and settlement data, as shown in Figure 33-71.

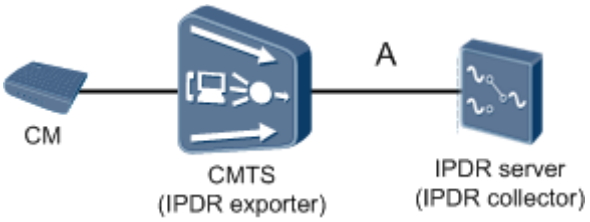
Figure 33-71 Basic IPDR network model



- Service consumer: a CM in the CMTS network
- Service element: a CMTS that collects IPDR data and communicates with an IPDR server. In this model, the CMTS, an MA5600T/MA5603T/MA5608T, is an IPDR exporter.
- Mediation system: an IPDR server that stores and forwards IPDR data. In this model, the IPDR server is an IPDR collector.
- Settlement data: an MSO back-end operating system, such as a business support system (BSS), an OSS, or a clearing house.

In Figure 33-71, the MA5600T/MA5603T/MA5608T collects IPDR data, encodes the collected data in XDR format, and sends the encoded data to the IPDR server through interface A. Then, the IPDR server stores and forwards the IPDR data to the MSO back-end operating system. Figure 33-72 shows role relationships.

Figure 33-72 Role relationships



Device Connection Mode

An IPDR server can work in active or passive mode. The MA5600T/MA5603T/MA5608T supports configuration of the connection mode. By default, the MA5600T/MA5603T/MA5608T works in active mode.

- In active mode, the MA5600T/MA5603T/MA5608T actively listens to the status of the socket port connected to the IPDR server. The IPDR server actively initiates a connection request to the MA5600T/MA5603T/MA5608T. After receiving the connection request, the MA5600T/MA5603T/MA5608T responds to the IPDR server. The connection is then set up.

- In passive mode, the IPDR server actively listens to the status of the port connected to the MA5600T/MA5603T/MA5608T. The MA5600T/MA5603T/MA5608T actively initiates a connection request to the IPDR server. After receiving the connection request, the IPDR server responds to the MA5600T/MA5603T/MA5608T. Then, the connection is set up.

Session

A session is a logical connection between the MA5600T/MA5603T/MA5608T and one or multiple IPDR servers. The session is used for delivering the collected data, including accounting information, CM running status, frequency spectrum information, debugging information, and CMTS statistics. Through a session, an IPDR server uses an IPDR statistics profile to collect and record data in a specified collection mode.

Multiple sessions, distinguished by session IDs, can be maintained concurrently in an IPDR exporter or collector.

ACK Timeout Duration

When initiating a session, the MA5600T/MA5603T/MA5608T sends a SESSION START message containing the ACK timeout duration to an IPDR server. If the IPDR server does not respond to the SESSION START message before the ACK timeout duration expires, the MA5600T/MA5603T/MA5608T fails to connect to the IPDR server.

A shorter ACK timeout duration results in a larger number of packets sent from the MA5600T/MA5603T/MA5608T that need to be acknowledged by the IPDR server, causing a heavy load on the IPDR server.

Keepalive Duration

The keepalive duration is used to detect an idle TCP/IP connection. The MA5600T/MA5603T/MA5608T sends a keepalive message to an IPDR server to check whether it can connect to the IPDR server. If the IPDR server does not respond within the keepalive duration, the MA5600T/MA5603T/MA5608T fails to connect to the IPDR server. Then, the MA5600T/MA5603T/MA5608T sends a DISCONNECT message to the IPDR server to terminate the session.

A shorter keepalive duration allows faster detection of IPDR server faults, but also results in a larger number of keepalive packets sent from the MA5600T/MA5603T/MA5608T to the IPDR server, burdening the IPDR server.

Maximum Number of Unacknowledged Messages

The maximum number of unacknowledged messages determines the maximum number of messages that the MA5600T/MA5603T/MA5608T can send to an IPDR server before receiving an ACK message from the IPDR server. When initiating a session, the MA5600T/MA5603T/MA5608T sends a SESSION START message containing the maximum number of unacknowledged messages to an IPDR server.

A larger number for maximum number of unacknowledged messages results in faster packet transmitting and receiving between the MA5600T/MA5603T/MA5608T and the IPDR server, but also results in a higher CPU usage on the IPDR server.

IPDR Statistics Profiles

The MA5600T/MA5603T/MA5608T uses the profiles defined by DOCSIS 3.0 to collect IPDR data in a specified format and data collection mode. The data collection mode varies depending on the IPDR statistics profile. Table 33-28 lists the IPDR statistics profiles and data collection modes.

Table 33-28 IPDR statistics profiles and data collection modes

DOCSIS 3.0 Profile	Description	Profile Name Defined by DOCSIS 3.0	Profile Name Defined by the MA5600T/MA5603T/MA5608T	Data Collection Mode
SAMIS	Accounting profile	SAMIS-TYPE1	SAMIS-TYPE 1	Periodic reporting or one-off reporting
		SAMIS-TYPE2	SAMIS-TYPE 2	Periodic reporting or one-off reporting
Diagnostic log	Diagnosis profile	DIAGLOG-TYPE	DIAGLOG-TYPE	One-off reporting
		DIAGLOG-EVENT-TYPE	DIAGLOG-EVENT-TYPE	Event-driven reporting
		DIAGLOG-DETAIL-TYPE	DIAGLOG-DETAIL-TYPE	Periodic reporting, one-off reporting, or event-driven reporting
Spectrum measurement	Frequency spectrum profile	SPECTRUM-MEASUREMENT-TYPE	SPECTRUM-MEASUREMENT	Periodic reporting or one-off reporting
CMTS CM registration status information	CM registration status profile	CMTS-CM-REG-STATUS-TYPE	CM-REG-STATUS	Periodic reporting, one-off reporting, or event-driven reporting
CMTS CM upstream status information	CM upstream service flow profile	CMTS-CM-US-STATS-TYPE	CM-US-STATS	Periodic reporting or one-off reporting
CMTS topology	CMTS topology profile	CMTS-TOPOLOGY-TYPE	TOPOLOGY	One-off reporting or event-driven reporting
CPE information	Customer premises equipment (CPE) profile	CPE-TYPE	CPE-TYPE	One-off reporting or event-driven reporting
CMTS	CMTS	CMTS-US-UTIL-	US-UTIL	Event-driven

DOCSIS 3.0 Profile	Description	Profile Name Defined by DOCSIS 3.0	Profile Name Defined by the MA5600T/MA5603T/MA5608T	Data Collection Mode
utilization statistics	bandwidth statistics profile	STATS-TYPE		reporting
		CMTS-DS-UTIL-STATS-TYPE	DS-UTIL	Event-driven reporting
CMTS service flow type	CMTS service flow statistics profile	DOCSIS-SERVICE-FLOW-TYPE	SERVICE-FLOW	One-off reporting or event-driven reporting

NOTE

- Select a DOCSIS 3.0 profile based on a site's requirements. For more details about DOCSIS 3.0 profiles, see DOCSIS 3.0 standards.
- The names listed in column **Profile Name Defined by the MA5600T/MA5603T/MA5608T** can be configured on the MA5600T/MA5603T/MA5608T.

IPDR Data Collection Modes

The MA5600T/MA5603T/MA5608T collects IPDR data in one of the following three modes: periodic reporting, one-off reporting, and event-driven reporting.

- Periodic reporting: The MA5600T/MA5603T/MA5608T sends a SESSION START message to an IPDR server, indicating the start of the data collection interval. Then, the MA5600T/MA5603T/MA5608T collects IPDR data, encodes the data in XDR format, and sends the data to the IPDR server in real time. After sending all IPDR data, the MA5600T/MA5603T/MA5608T sends a SESSION STOP message to the IPDR server, indicating the end of the data collection interval. The new period starts after the data collection interval elapses.
- One-off reporting: The data collection process between the MA5600T/MA5603T/MA5608T and the IPDR server in one-off reporting mode is the same as that in periodic reporting mode. In one-off reporting mode, however, the process occurs only once and ends when the MA5600T/MA5603T/MA5608T sends a SESSION STOP message to the IPDR server.
- Event-driven reporting: When an event occurs after a session is activated, the MA5600T/MA5603T/MA5608T reports an IPDR record to the IPDR server.

33.16.3 IPDR Networking Applications

The PON+D-CCAP solution can be used in centralized management mode or standalone NE mode. Table 33-29 lists the IPDR networking differences between the two modes.

Table 33-29 IPDR networking differences between centralized management mode and standalone NE mode

Networking	NE Management	Data Transfer Process

Networking	NE Management	Data Transfer Process
Centralized management mode	The optical line terminal (OLT) and the MA5633 constitute a CMTS. The OLT is managed by the U2000.	The IPDR server obtains IPDR data from the OLT. Figure 33-73 shows the IPDR networking in centralized management mode.
Standalone NE mode	The OLT and the MA5633 are separately managed by the U2000.	The IPDR server obtains IPDR data from the MA5633. Figure 33-74 shows the IPDR networking in standalone NE mode.

Figure 33-73 IPDR networking in centralized management mode

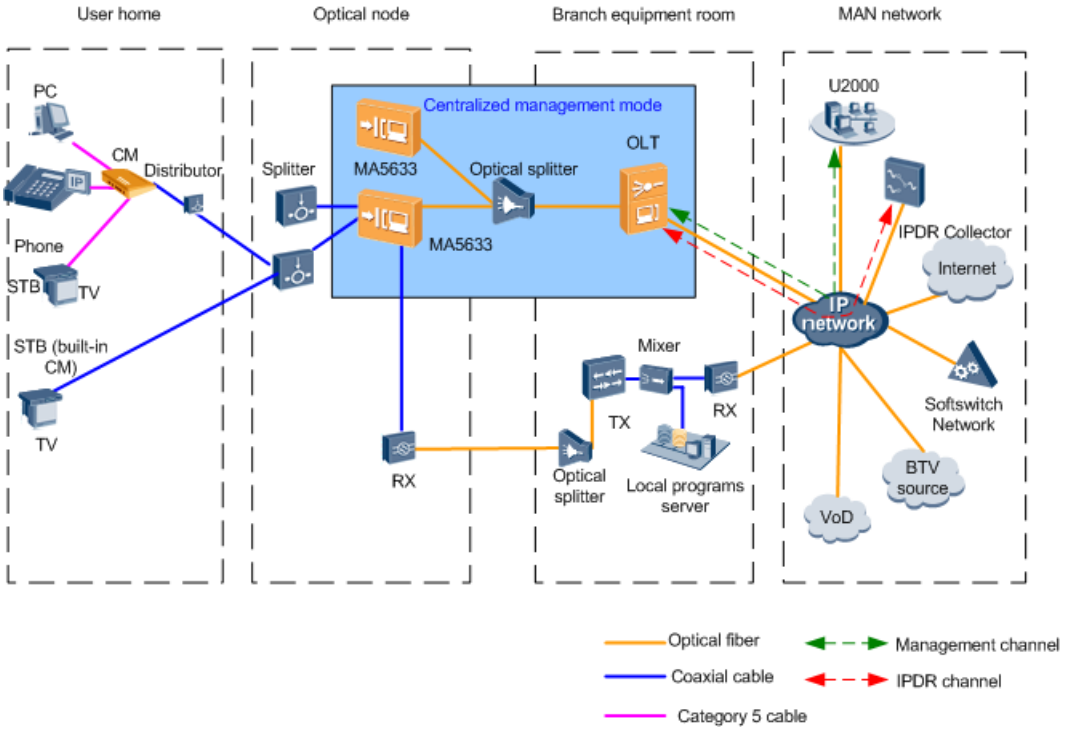
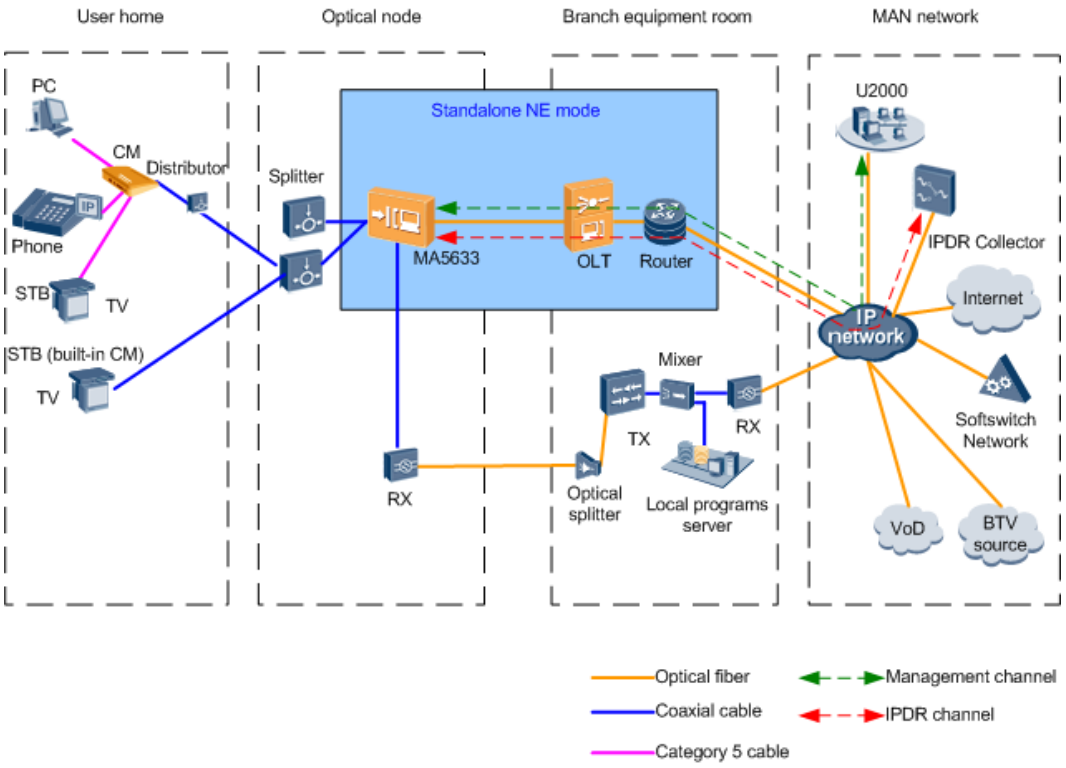


Figure 33-74 IPDR networking in standalone NE mode



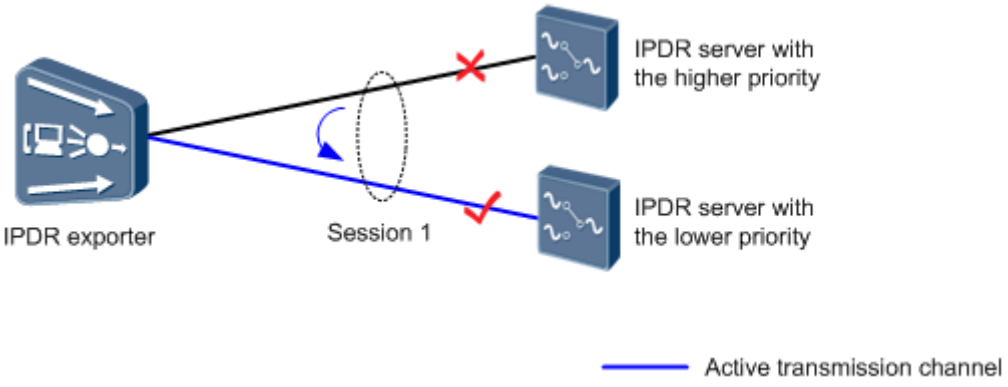
33.16.4 IPDR Server Protection Switchover

The IPDR feature allows an IPDR session to be associated with two IPDR servers. The IPDR server with the higher priority is used as the active server. The IPDR server with the lower priority is used as the standby server. If the IPDR server with the higher priority fails, the MA5600T/MA5603T/MA5608T automatically switches to the lower-priority IPDR server. This helps prevent key data from being lost during potential link disconnections.

Switchover Process

During the switchover process, the MA5600T/MA5603T/MA5608T sends only the IPDR data that has not been acknowledged by the higher-priority IPDR server to the lower-priority server. Figure 33-75 shows the IPDR server switchover process.

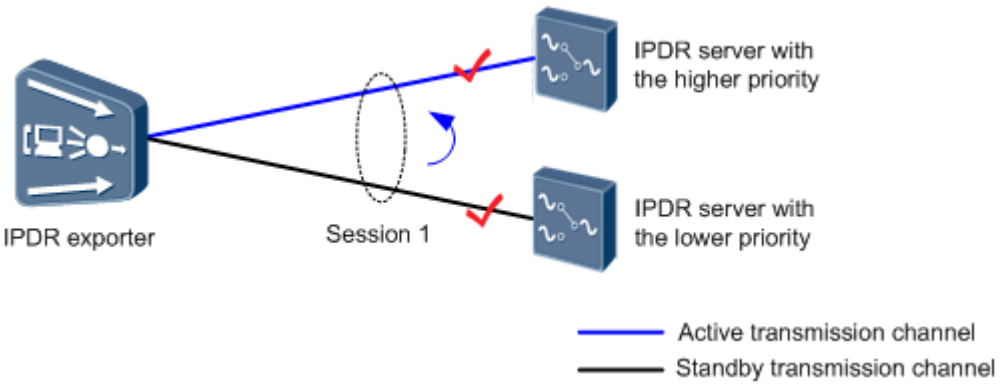
Figure 33-75 IPDR server switchover process



Switchover Restoration Process

After the MA5600T/MA5603T/MA5608T switches to the IPDR server with the lower priority, it will still attempt to communicate with the higher-priority IPDR server. After the higher-priority IPDR server recovers, the MA5600T/MA5603T/MA5608T switches back to this IPDR server. Figure 33-76 shows the IPDR server switchover restoration process.

Figure 33-76 IPDR server switchover restoration process



33.16.5 Configuring IPDR

This section describes how to configure the IPDR feature on the MA5600T/MA5603T/MA5608T through the CLI. After the configuration, the MA5600T/MA5603T/MA5608T can communicate with an IPDR server, collect IPDR data, and send the data to the IPDR server.

Prerequisites

- The IPDR server is functioning properly.
- The MA5600T/MA5603T/MA5608T can communicate with the IPDR server.

Configuration Process

Figure 33-77 shows the flowchart for configuring the IPDR feature.

Figure 33-77 Flowchart for configuring the IPDR feature

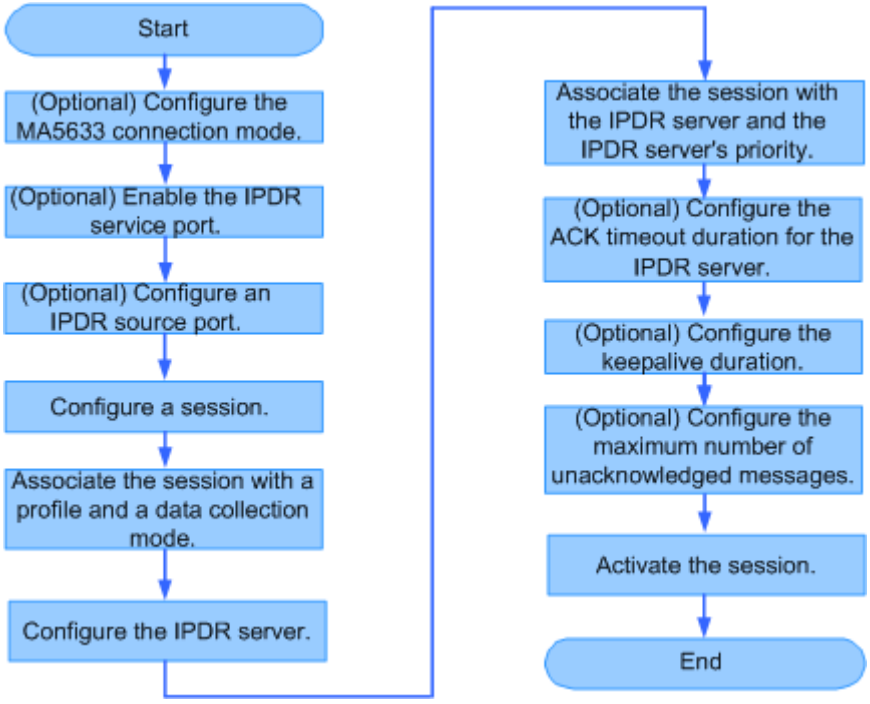


Table 33-30 lists and describes the commands corresponding to the steps in Figure 33-77.

Table 33-30 Procedure for configuring the IPDR feature

No.	Step	Command	Remarks
1	(Optional) Configure the MA5600T/MA5603T/MA5608T connection mode.	**ipdr collector type** { **active** | **passive** }	The MA5600T/MA5603T/MA5608T connects to an IPDR server in active mode by default. If the MA5600T/MA5603T/MA5608T exchanges data with an IPDR server in passive mode, change the connection mode to **passive**.
2	(Optional) Enable the IPDR service port.	**sysman service ipdr enable**	The IPDR service port is enabled by default. Perform this step if the IPDR service port is disabled. The MA5600T/MA5603T/MA5608T can communicate with an IPDR server only after the IPDR service port is enabled.
3	(Optional) Configure an IPDR source port.	**sysman source**	By default, IPDR data is transmitted through the source port specified by a routing protocol. You can specify the source port through which the IPDR data is transmitted.
4	Configure a session.	**ipdr session** *id* **name** *name*	The session name and session description can be modified. A session can be deleted only after it is deactivated. You

No.	Step	Command	Remarks
			can run the **ipdr session deactivate** command to deactivate a session. **NOTE** Deactivated sessions can no longer be used.
5	Associate the session with a profile and data collection mode.	**ipdr session** *id* **template** *template-name* **type** { **ad-hoc** | **event** | **interval** *interval* }	A profile and data collection mode can be associated with only one session.
6	Configure the IPDR server.	**ipdr collector name** *name* **ip** *ip-address* [**port** *port*]	• A configured IPDR server cannot be modified. • An IPDR server has a unique server name and IP address. • The default port number of the IPDR server is 4737. When the MA5600T/MA5603T/MA5608T connects to an IPDR server in active mode, you do not need to configure the port number.
7	Associate the session with the IPDR server and the IPDR server's priority.	**ipdr session** *id* **collector** *name* **priority** *priority*	Session configurations cannot be modified if the session is activated. To modify the configuration, you must first deactivate the session.
8	(Optional) Configure the ACK timeout duration for the IPDR server.	**ipdr parameter ack-timeout**	The value of this parameter can be changed only after all sessions associated with the IPDR server are deactivated.
9	(Optional) Configure the keepalive duration.	**ipdr parameter keepalive**	The value of this parameter can be changed only after all sessions associated with the IPDR server are deactivated.
10	(Optional) Configure the maximum number of unacknowledged messages.	**ipdr parameter max-unacked**	The value of this parameter can be changed only after all sessions associated with the IPDR server are deactivated.
11	Activate the session.	**ipdr session activate** { **all** | *id* }	A session can be activated only after the profile, data collection mode, and IPDR server associated with the session are

No.	Step	Command	Remarks
			configured.

Configuration Example and Data Planning

This section describes how to configure the IPDR feature in standalone NE mode. The configuration method in centralized management mode is the same as that in standalone NE mode. The only difference is that the configuration is performed on an OLT in centralized mode and on the MA5633 in standalone NE mode.

Configuration Object	Data
MA5600T/MA5603T/MA5608T	Connection mode: active
Session	• ID: 6 • Name: huawei • Description: huawei1
Data collection	• IPDR profile: SAMIS-TYPE1 • Data collection mode: periodic reporting; interval: 15 minutes
IPDR server	• Name: collector_server1 • IP address: 10.10.10.10 • Port number: 1 • Priority: 1

Procedure

Configure a session.

```
huawei(config)#ipdr session 6 name huawei desc huawei1
```

Step 1 Associate the session with a profile type and specify the data collection mode.

```
huawei(config)#ipdr session 6 template SAMIS-TYPE1 type interval 15
```

Step 2 Configure the IPDR server.

```
huawei(config)#ipdr collector name collector_server1 ip 10.10.10.10 port 1
```

Step 3 Associate the session with the IPDR server and the IPDR server's priority.

```
huawei(config)#ipdr session 6 collector collector_server1 priority 1
```

Step 4 Activate the session.

```
huawei(config)#ipdr session activate 6
```

----End

Result

After the configuration, the MA5600T/MA5603T/MA5608T collects IPDR data and sends the data to the IPDR server. The IPDR server receives all the IPDR data repeatedly every 15 minutes.

Follow-up Procedure

The OSS or BSS analyzes the IPDR server data.

33.16.6 IPDR Reference Files

IPDR reference files are as follows:

- DOCSIS 2.0, DOCSIS 3.0, Euro-DOCSIS 2.0, and Euro-DOCSIS 3.0
- TMF8000-IPDR-IIS-PS from TM Forum
- TMF8001-IPDR-IIS-PS from TM Forum

33.17 PNM

As cable networks evolve, many diverse services are carried over them. Accordingly, operators are eager to take effective measures in O&M, such as monitoring services and diagnosing faults, to fix problems before they have any impact on services. Proactive network maintenance (PNM) enables a cable modem termination system (CMTS) to detect a fault on the network so that the CMTS can actively identify and rectify the fault before this fault adversely affects user services, thereby reducing network O&M costs and improving operators' service level agreement (SLA).

What Is PNM

Traditional Reactive Network Maintenance

In traditional reactive network maintenance, a fault is located and rectified only after it adversely affects user services. In addition, the fault location is identified by maintenance personnel based on their experience, or is identified aimlessly even. Common reactive network maintenance scenarios are as follows:

- After user experience is deteriorated, for example, the Internet access rate becomes slow or TV images become unclear, the user reports the fault to the customer service center. Then, the maintenance personnel locate and rectify this fault.
- The U2000 monitors line running indicators in real time and reports an alarm when detecting an indicator exception, for example, severe bit errors are detected in forward error correction (FEC) statistics, the level is excessively low, or the signal-to-noise ratio (SNR) is small. Then, the maintenance personnel rectify the fault.

Proactive Network Maintenance

PNM relies on pre-equalization coefficients. Specifically, the CMTS monitors the pre-equalization coefficient of each cable modem (CM) on the hybrid fiber coaxial (HFC). By analyzing these pre-equalization coefficients, the CMTS identifies slight line running indicator changes and detects a latent fault. Based on the analysis results and CM topology,

the maintenance personnel locate and rectify the fault before the fault adversely affects user services.

A pre-equalization coefficient is used to compensate line distortion caused by mismatched impedance, an insecurely connected connector, or a damaged cable. Therefore, line fault information can be identified in a pre-equalization coefficient. A pre-equalization coefficient reflects a fault location while providing the fault type and severity.

33.17.2 Pre-equalization

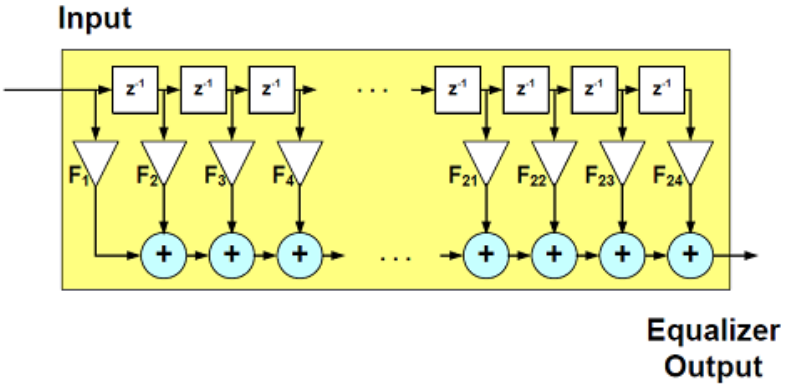
On HFC networks, signal impairments can be classified as linear impairments and nonlinear impairments. In a linear impairment, a signal will be changed in amplitude and phase compared with the original signal. In a nonlinear impairment, a signal will generate harmonics. Linear impairments involve group delay and micro-reflection.

- A group delay is determined based on active amplifiers on the HFC network. A greater number of amplifiers lead to a greater group delay.
- A micro-reflection is a signal reflection caused by mismatched line impedance.

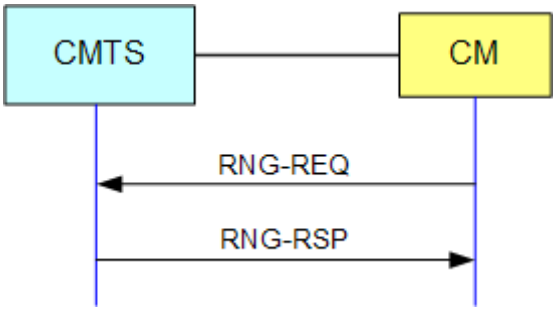
The pre-equalization technology minimizes impact of the two types of linear impairments on network data transmission.

Pre-equalization Working Principles

After a CM sends signals upstream, these signals pass through 24 delay taps. Each tap provides an amplitude. Then, the output signals are pre-equalized in time domains. In this way, signal transmission unflatness can be eliminated in frequency domains, and CMTS receives the signals with linear impairment canceled.



A CMTS obtains the pre-equalization coefficients of the CMs connected to it through ranging. The following figure shows the ranging exchange process.



After receiving the upstream RNG-REQ (ranging request) message sent from a CM, the CMTS uses an algorithm to calculate the pre-equalization coefficient for the CM. Then, the

CMTS uses an RNG-RSP (ranging response) message to notify the CM of this pre-equalization coefficient for upstream data transmission. The pre-equalization effect may fail to meet the CMTS's requirements on receiving data after one-round of ranging exchange between the CMTS and CM. Therefore, the CMTS and the CM must exchange data multiple times for a stable pre-equalization coefficient. The ranging lasts during the data exchanging. After the pre-equalization coefficient becomes stable, the CMTS still continuously checks whether this pre-equalization coefficient is required to adjust.

33.17.3 Process of Locating an HFC Network Fault Using PNM

Pre-equalization PNM-based HFC network fault locating involves the following processes:

Data Collection

Collects pre-equalization coefficients and other parameters assisting fault locating, such as frequency, frequency bandwidth, and SNR.

Data Analysis

Uses a pre-equalization PNM-based algorithm to analyze data.

The U2000 collects data on the OLT and CMTSs. The smart service assurance (SSA) system periodically analyzes the collected data.

Severity Evaluation

Provides a micro-reflection level for each CM based on PNM data analysis results. The maintenance personnel only need to focus on the CMs with a high micro-reflection level and check whether the coaxial lines connected to such CMs are faulty.

Micro-reflection levels can be green, yellow, or red, which are in ascending order in severity.

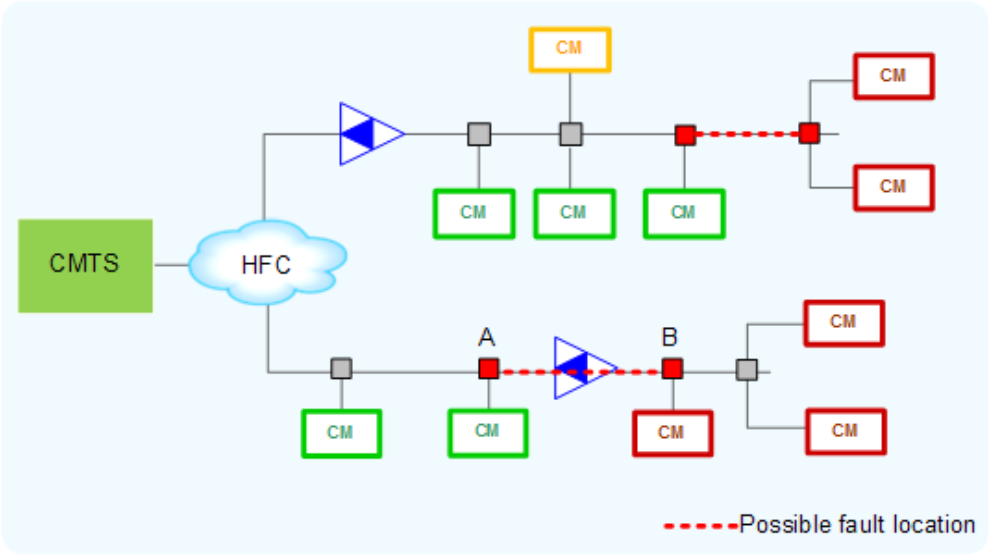
- Green indicates normal and that no action is necessary.
- Yellow indicates that the CM performance has been sightly deteriorated but the fault has not adversely affected user services. If this happens, the CM should be monitored more frequently.
- Red implies the need for immediate action because the fault may have adversely affected user services.

Fault Locating

Each CM functions as a fault detection probe in the PNM system. The CMTS monitors the linear impairments of each CM in real time. Then, the maintenance personnel can determine whether a network fault is related only to one CM or multiple CMs based on CM topology. If

the network fault is related only to one CM, the maintenance personnel are required to check only the coaxial line connected to the CM. If the network fault is related to multiple CMs, the maintenance personnel are required to check the coaxial line shared between these CMs.

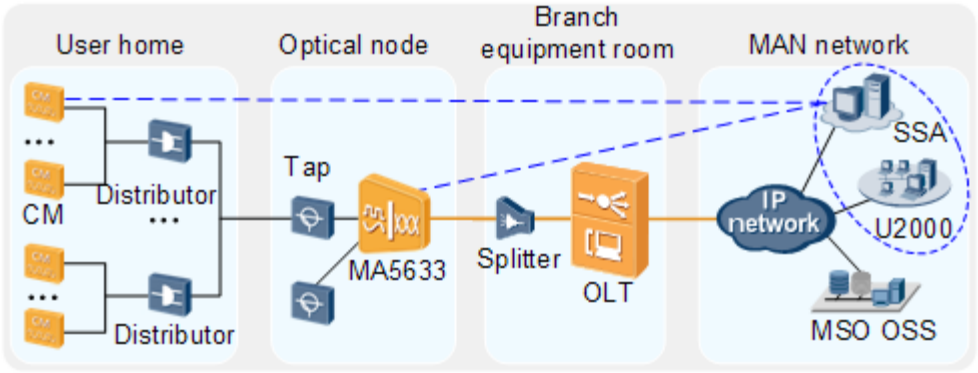
In the following figure, the micro-reflection levels of the three CMs in the lower-level of point B are high (red), and those of the CMs in the upper-level of point B are normal (green or yellow). Therefore, the fault must occur on the coaxial line between points A and B, requiring the maintenance personnel to locate the fault onsite.



33.17.4 Application Scenarios

PNM applies to the two scenarios: OLT+MA5633 in standalone NE mode and OLT+MA5633 in centralized management mode.

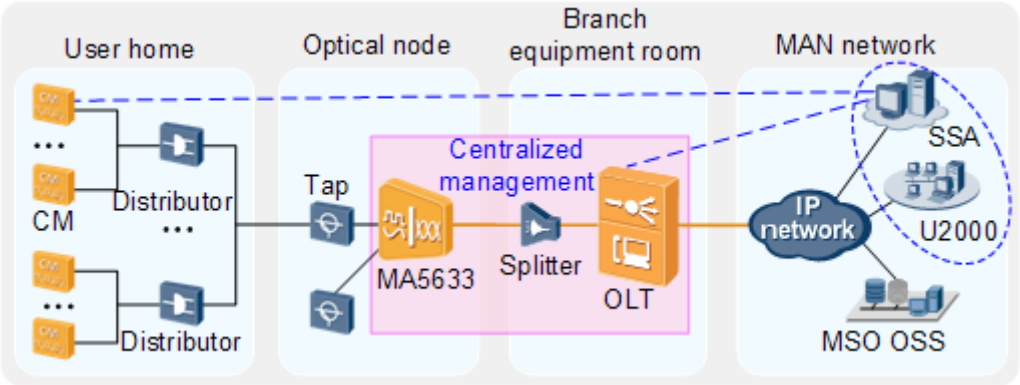
Standalone NE Mode



In the preceding figure,

- The MA5633 connects to the OLT or switch using GPON or GE upstream transmission. Then, the OLT transmits data upstream to the IP network using a router or Layer 3 switch. In this network, the OLT works in Layer 3 forwarding mode.
- The service provisioning system, management system, U2000, and SSA are deployed on the head end. SSA functions as the PNM server, which uses the network management interface to collect PNM data from the MA5633, process the data, and display processing results.

Centralized Management Mode



In the preceding figure,

- The MA5633 connects to the OLT or switch using GPON or GE upstream transmission. Then, the OLT transmits data upstream to the IP network using a router or Layer 3 switch. In this network, the OLT works in Layer 3 forwarding mode.
- The service provisioning system, management system, U2000, and SSA are deployed on the head end. SSA functions as the PNM server, which uses the network management interface to collect PNM data from the OLT, process the data, and display processing results.

33.17.5 PNM Functions

PNM data is processed and displayed on the PNM server (U2000). This section describes PNM functions.

DOCSIS 3.0-compliant PNM

Monitors the following line parameters:

- Constellation diagrams before and after pre-equalization: Identifies and locates active component issues and noises. This scenario features large-volume data and strict real-time requirements.
- Noises in channel-based CM carriers: Provides noise spectrum graphs and noise levels in an in-service channel for separating noises from signals.
- MERs before and after equalization: Measures the MERs, especially those before equalization for determining the quality of a cable.
- Periodically collected pulse noises: Detects short (a few milliseconds) pulse noises that are difficult to identify.

DOCSIS 3.1-compliant PNM

- Spectrum analysis: Scans spectra.
- Vector signal analysis: Collects equalization coefficients, evaluates downstream channels, displays constellation diagrams, and measures MERs based on subcarriers.
- Network analysis: Collects FEC and pulse noise statistics.

33.17.6 Standards and Protocols Compliance

Proactive Network Maintenance Using Pre-equalization