QTECH

МИР ДОСТУПНЕЕ

QSW-8300 series

CONFIGURATION MANUAL

Content

+7(495) 797-3311 www.qtech.ru
Москва, Новозаводская ул., 18, стр. 1

Chapter 1 Port Configuration

1.1 Introduction to Port

Switch contains Cable ports and Combo ports. The Combo ports can be configured as either 1000GX-TX ports or SFP Gigabit fiber ports.

If the user needs to configure some network ports, he/she can use the interface ethernet <interface-list> command to enter the appropriate Ethernet port configuration mode, where <interface-list> stands for one or more ports. If <interface-list> contains multiple ports, special characters such as ';' or '-' can be used to separate ports, ';' is used for discrete port numbers and '-' is used for consecutive port numbers. Suppose an operation should be performed on ports 2, 3, 4, 5, the command would look like: interface ethernet 1/0/2-5. Port speed, duplex mode and traffic control can be configured under Ethernet Port Mode causing the performance of the corresponding network ports to change accordingly.

1.2 Network Port Configuration Task List

1. Enter the network port configuration mode
2. Configure the properties for the network ports
(1) Configure combo mode for combo ports
(2) Enable/Disable ports
(3) Configure port names
(4) Configure port cable types
(5) Configure port speed and duplex mode
(6) Configure bandwidth control
(7) Configure traffic control
(8) Enable/Disable port loopback function
(9) Configure broadcast storm control function for the switch
(10) Configure scan port mode
(11) Configure rate-violation control of the port
(12) Configure interval of port-rate-statistics
3. Virtual cable test

1. Enter the Ethernet port configuration mode

Command	Explanation
Global Mode	
interface ethernet *<interface-list>*	Enters the network port configuration mode.

2. Configure the properties for the Ethernet ports

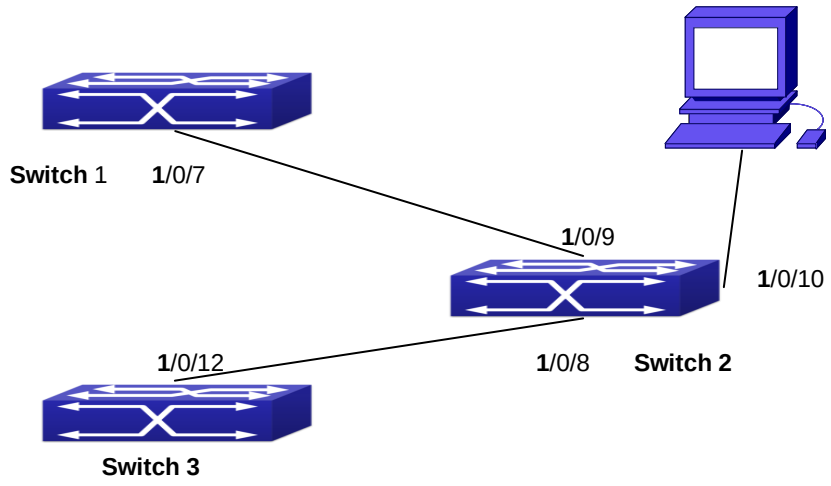
Command	Explanation
Port Mode	
combo-forced-mode {copper-forced | sfp-forced}	Sets the combo port mode (combo ports only).
shutdown **no shutdown**	Enables/Disables specified ports.
name <string> **no name**	Names or cancels the name of specified ports.
mdi { auto | across | normal } **no mdi**	Sets the cable type for the specified port; this command is not supported by combo port and fiber port of switch.
speed-duplex {auto [10 [100 [1000]] [auto | full | half |]] | force10-half | force10-full | force100-half | force100-full | force100-fx [module-type {auto-detected | no-phy-integrated | phy-integrated}] | {{force1g-half | force1g-full} [nonegotiate [master | slave]]}| force10g-full} **no speed-duplex**	Sets port speed and duplex mode of 100/1000Base-TX or 100Base-FX ports. The no format of this command restores the default setting, i.e., negotiates speed and duplex mode automatically.
negotiation {on|off}	Enables/Disables the auto-negotiation function of 1000Base-FX ports.
bandwidth control <*bandwidth*> [both | receive | transmit] **no bandwidth control**	Sets or cancels the bandwidth used for incoming/outgoing traffic for specified ports.
flow control **no flow control**	Enables/Disables traffic control function for specified ports.
loopback **no loopback**	Enables/Disables loopback test function for specified ports.
rate-suppression {dlf | broadcast | multicast} <packets>	Enables the storm control function for broadcasts, multicasts and unicasts with unknown destinations (short for broadcast), and sets the allowed broadcast packet number; the no format of this command disables the broadcast storm control function.
port-scan-mode {interrupt | poll} **no port-scan-mode**	Configure port-scan-mode as interrupt or poll mode, the no command restores the default port-scan-mode.
rate-violation <200-2000000> [recovery	Set the max packet reception rate of a port. If

<0-86400>|] **no rate-violation**	the rate of the received packet violates the packet reception rate, shut down this port and configure the recovery time, the default is 300s. The no command will disable the rate-violation function of a port.
Global Mode	
port-rate-statistics interval [<interval -value>]	Configure the interval of port-rate-statistics.

3. Virtual cable test

Command	Explanation
Port Configuration Mode	
virtual-cable-test	Test virtual cables of the port.

1.3 Port Configuration Example



Port Configuration Example

No VLAN has been configured in the switches, default VLAN1 is used.

Switch	Port	Property
Switch1	1/0/7	Ingress bandwidth limit: 50 M
Switch2	1/0/8	Mirror source port
	1/0/9	100Mbps full, mirror source port
	1/0/10	1000Mbps full, mirror destination port
Switch3	1/0/12	100Mbps full

The configurations are listed below:

Switch1:

Switch1(config)#interface ethernet 1/0/7
Switch1(Config-If-Ethernet1/0/7)#bandwidth control 50 both
Switch2:
Switch2(config)#interface ethernet 1/0/9
Switch2(Config-If-Ethernet1/0/9)#speed-duplex force100-full
Switch2(Config-If-Ethernet1/0/9)#exit
Switch2(config)#interface ethernet 1/0/10
Switch2(Config-If-Ethernet1/0/10)#speed-duplex force1g-full
Switch2(Config-If-Ethernet1/0/10)#exit
Switch2(config)#monitor session 1 source interface ethernet1/0/8;1/0/9
Switch2(config)#monitor session 1 destination interface ethernet 1/0/10
Switch3:
Switch3(config)#interface ethernet 1/0/12
Switch3(Config-If-Ethernet1/0/12)#speed-duplex force100-full
Switch3(Config-If-Ethernet1/0/12)#exit

1.4 Port Troubleshooting

Here are some situations that frequently occurs in port configuration and the advised solutions:
Two connected fiber interfaces won't link up if one interface is set to auto-negotiation but the other to forced speed/duplex. This is determined by IEEE 802.3.
The following combinations are not recommended: enabling traffic control as well as setting multicast limiting for the same port; setting broadcast, multicast and unknown destination unicast control as well as port bandwidth limiting for the same port. If such combinations are set, the port throughput may fall below the expected performance.
For Combo port, it supports copper-forced mode and sfp-forced mode (default mode), here, copper port will not be up.

Chapter 2 Port Isolation Function Configuration

2.1 Introduction to Port Isolation Function

Port isolation is an independent port-based function working in an inter-port way, which isolates flows of different ports from each other. With the help of port isolation, users can isolate ports within a VLAN to save VLAN resources and enhance network security. After this function is configured, the ports in a port isolation group will be isolated from each other, while ports belonging to different isolation groups or no such group can forward data to one another normally. No more than 16 port isolation groups can a switch have.

2.2 Task Sequence of Port Isolation

1. Create an isolate port group
2. Add Ethernet ports into the group
3. Specify the flow to be isolated
4. Display the configuration of port isolation

1. Create an isolate port group

Command	Explanation
Global Mode	
isolate-port group <*WORD*> **no isolate-port group <*WORD*>**	Set a port isolation group; the no operation of this command will delete the port isolation group.

2. Add Ethernet ports into the group

Command	Explanation
Global Mode	
isolate-port group <*WORD*> switchport interface [ethernet | port-channel] <*IFNAME*> **no isolate-port group <*WORD*> switchport interface [ethernet | port-channel] <*IFNAME*>**	Add one port or a group of ports into a port isolation group to isolate, which will become isolated from the other ports in the group; the no operation of this command will remove one port or a group of ports out of a port isolation group.

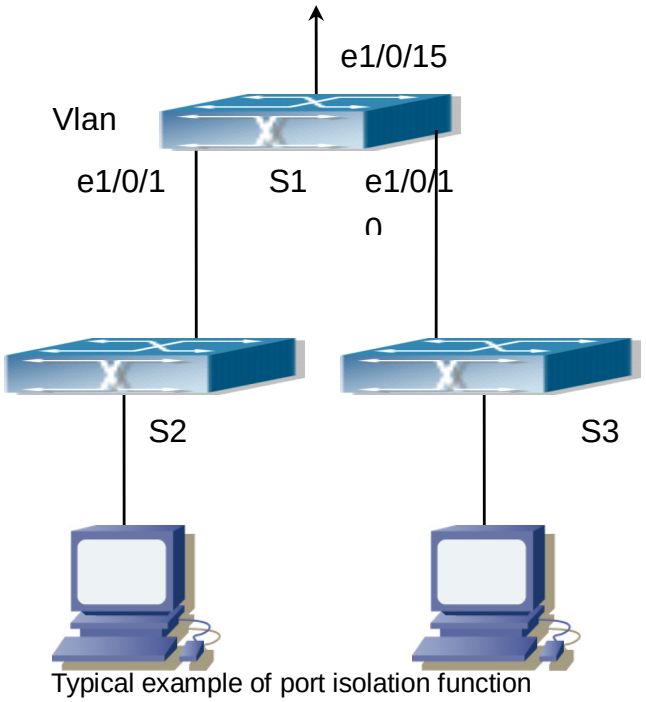
3. Specify the flow to be isolated

Command	Explanation
Global Mode	
isolate-port apply [<l2|l3|all>]	Apply the port isolation configuration to isolate layer-2 flows, layer-3 flows or all flows.

4. Display the configuration of port isolation

Command	Explanation
Admin Mode and global Mode	
show isolate-port group [*<WORD>*]	Display the configuration of port isolation, including all configured port isolation groups and Ethernet ports in each group.

2.3 Port Isolation Function Typical Examples



Typical example of port isolation function

The topology and configuration of switches are showed in the figure above, with e1/0/1, e1/0/10 and e1/0/15 all belonging to VLAN 100. The requirement is that, after port isolation is enabled on switch S1, e1/0/1 and e1/0/10 on switch S1 can not communicate with each other, while both of them can communicate with the uplink port e1/0/15. That is, the communication between any pair of downlink ports is disabled while that between any downlink port and a specified uplink port is normal. The uplink port can communicate with any port normally.

The configuration of S1:

Switch(config)#isolate-port group test
Switch(config)#isolate-port group test switchport interface ethernet 1/0/1;1/0/10

Chapter 3 Port Loopback Detection Function Configuration

3.1 Introduction to Port Loopback Detection Function

With the development of switches, more and more users begin to access the network through Ethernet switches. In enterprise network, users access the network through layer-2 switches, which means urgent demands for both internet and the internal layer 2 Interworking. When layer 2 Interworking is required, the messages will be forwarded through MAC addressing the accuracy of which is the key to a correct Interworking between users. In layer 2 switching, the messages are forwarded through MAC addressing. Layer 2 devices learn MAC addresses via learning source MAC address, that is, when the port receives a message from an unknown source MAC address, it will add this MAC to the receive port, so that the following messages with a destination of this MAC can be forwarded directly, which also means learn the MAC address once and for all to forward messages.

When a new source MAC is already learnt by the layer 2 device, only with a different source port, the original source port will be modified to the new one, which means to correspond the original MAC address with the new port. As a result, if there is any loopback existing in the link, all MAC addresses within the whole layer 2 network will be corresponded with the port where the loopback appears (usually the MAC address will be frequently shifted from one port to another), causing the layer 2 network collapsed. That is why it is a necessity to check port loopbacks in the network. When a loopback is detected, the detecting device should send alarms to the network management system, ensuring the network manager is able to discover, locate and solve the problem in the network and protect users from a long-lasting disconnected network.

Since detecting loopbacks can make dynamic judgment of the existence of loopbacks in the link and tell whether it has gone, the devices supporting port control (such as port isolation and port MAC address learning control) can maintain that automatically, which will not only reduce the burden of network managers but also response time, minimizing the effect caused loopbacks to the network.

3.2 Port Loopback Detection Function Configuration Task List

Configure the time interval of loopback detection
Enable the function of port loopback detection
Configure the control method of port loopback detection
Display and debug the relevant information of port loopback detection
Configure the loopback-detection control mode (automatic recovery enabled or not)

1. Configure the time interval of loopback detection

Command	Explanation
Global Mode	
loopback-detection interval-time *<loopback> <no-loopback>* **no loopback-detection interval-time**	Configure the time interval of loopback detection.

2. Enable the function of port loopback detection

Command	Explanation
Port Mode	
loopback-detection specified-vlan <vlan-*list*> **no loopback-detection specified-vlan <vlan-*list*>**	Enable and disable the function of port loopback detection.

3. Configure the control method of port loopback detection

Command	Explanation
Port Mode	
loopback-detection control {shutdown |block| learning} **no loopback-detection control**	Enable and disable the function of port loopback detection control.

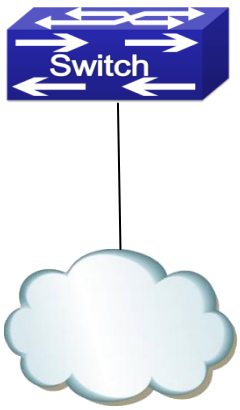
4. Display and debug the relevant information of port loopback detection

Command	Explanation
Admin Mode	
debug loopback-detection **no debug loopback-detection**	Enable the debug information of the function module of port loopback detection. The no operation of this command will disable the debug information.
show loopback-detection [interface *<interface-list>*]	Display the state and result of the loopback detection of all ports, if no parameter is provided; otherwise, display the state and result of the corresponding ports.

5. Configure the loopback-detection control mode (automatic recovery enabled or not)

Command	Explanation
Global Mode	
loopback-detection control-recovery timeout <0-3600>	Configure the loopback-detection control mode (automatic recovery enabled or not) or recovery time.

3.3 Port Loopback Detection Function Example



Typical example of port loopback detection

As shown in the above configuration, the switch will detect the existence of loopbacks in the network topology. After enabling the function of loopback detection on the port connecting the switch with the outside network, the switch will notify the connected network about the existence of a loopback, and control the port on the switch to guarantee the normal operation of the whole network.

The configuration task sequence of switch:

Switch(config)#loopback-detection interval-time 35 15

Switch(config)#interface ethernet 1/0/1

Switch(Config-If-Ethernet1/0/1)#loopback-detection special-vlan 1-3

Switch(Config-If-Ethernet1/0/1)#loopback-detection control block

If adopting the control method of block, MSTP should be globally enabled. And the corresponding relation between the spanning tree instance and the VLAN should be configured.

Switch(config)#spanning-tree

Switch(config)#spanning-tree mst configuration

Switch(Config-Mstp-Region)#instance 1 vlan 1

Switch(Config-Mstp-Region)#instance 2 vlan 2

Switch(Config-Mstp-Region)#

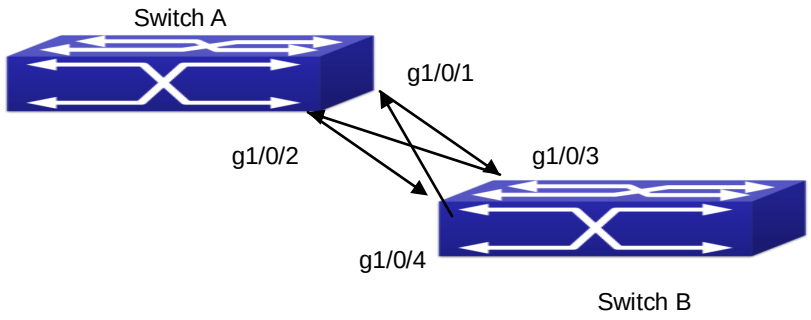
3.4 Port Loopback Detection Troubleshooting

The function of port loopback detection is disabled by default and should only be enabled if required.

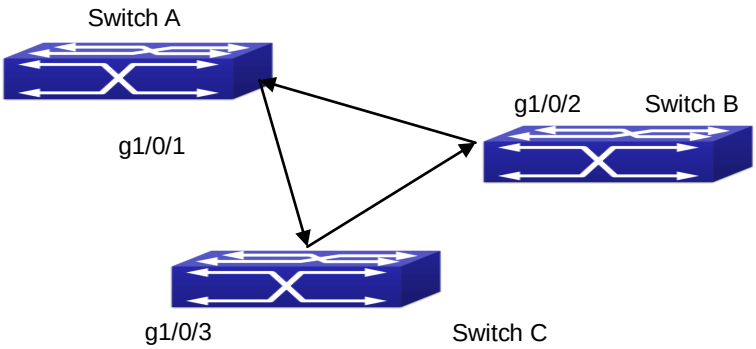
Chapter 4 ULDP Function Configuration

4.1 Introduction to ULDP Function

Unidirectional link is a common error state of link in networks, especially in fiber links. Unidirectional link means that only one port of the link can receive messages from the other port, while the latter one can not receive messages from the former one. Since the physical layer of the link is connected and works normal, via the checking mechanism of the physical layer, communication problems between the devices can not be found. As shown in Graph, the problem in fiber connection can not be found through mechanisms in physical layer like automatic negotiation.



Fiber Cross Connection



One End of Each Fiber Not Connected

This kind of problem often appears in the following situations: GBIC (Giga Bitrate Interface

Converter) or interfaces have problems, software problems, hardware becomes unavailable or operates abnormally. Unidirectional link will cause a series of problems, such as spinning tree topological loop, broadcast black hole.

ULDP (Unidirectional Link Detection Protocol) can help avoid disasters that could happen in the situations mentioned above. In a switch connected via fibers or copper Ethernet line (like ultra five-kind twisted pair), ULDP can monitor the link state of physical links. Whenever a unidirectional link is discovered, it will send warnings to users and can disable the port automatically or manually according to users' configuration.

The ULDP of switches recognizes remote devices and check the correctness of link connections via interacting ULDP messages. When ULDP is enabled on a port, protocol state machine will be started, which means different types of messages will be sent at different states of the state machine to check the connection state of the link by exchanging information with remote devices. ULDP can dynamically study the interval at which the remote device sends notification messages and adjust the local TTL (time to live) according to that interval. Besides, ULDP provides the reset mechanism, when the port is disabled by ULDP, it can check again through reset mechanism. The time intervals of notification messages and reset in ULDP can be configured by users, so that ULDP can respond faster to connection errors in different network environments.

The premise of ULDP working normally is that link works in duplex mode, which means ULDP is enabled on both ends of the link, using the same method of authentication and password.

4.2 ULDP Configuration Task Sequence

1. Enable ULDP function globally
2. Enable ULDP function on a port
3. Configure aggressive mode globally
4. Configure aggressive mode on a port
5. Configure the method to shut down unidirectional link
6. Configure the interval of Hello messages
7. Configure the interval of Recovery
8. Reset the port shut down by ULDP
9. Display and debug the relative information of ULDP

1. Enable ULDP function globally

Command	Explanation
Global configuration mode	
uldp enable **uldp disable**	Globally enable or disable ULDP function.

2. Enable ULDP function on a port

Command	Explanation
Port configuration mode	
uldp enable **uldp disable**	Enable or disable ULDP function on a port.

3. Configure aggressive mode globally

Command	Explanation
Global configuration mode	
uldp aggressive-mode **no uldp aggressive-mode**	Set the global working mode.

4. Configure aggressive mode on a port

Command	Explanation
Port configuration mode	
uldp aggressive-mode **no uldp aggressive-mode**	Set the working mode of the port.

5. Configure the method to shut down unidirectional link

Command	Explanation
Global configuration mode	
uldp manual-shutdown **no uldp manual-shutdown**	Configure the method to shut down unidirectional link.

6. Configure the interval of Hello messages

Command	Explanation
Global configuration mode	
uldp hello-interval <*integer*> **no uldp hello-interval**	Configure the interval of Hello messages, ranging from 5 to 100 seconds. The value is 10 seconds by default.

7. Configure the interval of Recovery

Command	Explanation
Global configuration mode	
uldp recovery-time <*integer*> **no uldp recovery-time <*integer*>**	Configure the interval of Recovery reset, ranging from 30 to 86400 seconds. The value is 0 second by default.

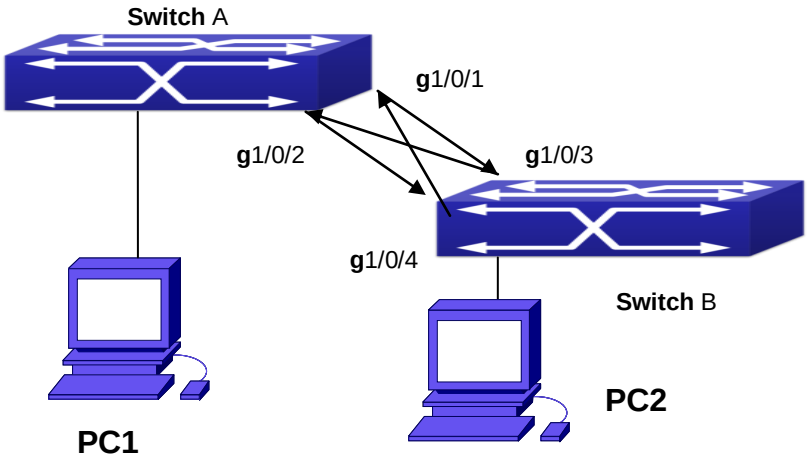
8. Reset the port shut down by ULDP

Command	Explanation
Global configuration mode or port configuration mode	
uldp reset	Reset all ports in global configuration mode; Reset the specified port in port configuration mode.

9. Display and debug the relative information of ULDP

Command	Explanation
Admin mode	
show uldp [interface ethernet IFNAME]	Display ULDP information. No parameter means to display global ULDP information. The parameter specifying a port will display global information and the neighbor information of the port.
debug uldp fsm interface ethernet <*IFname*> **no debug uldp fsm interface ethernet <*IFname*>**	Enable or disable the debug switch of the state machine transition information on the specified port.
debug uldp error **no debug uldp error**	Enable or disable the debug switch of error information.
debug uldp event **no debug uldp event**	Enable or disable the debug switch of event information.
debug uldp packet {receive|send} **no debug uldp packet {receive|send}**	Enable or disable the type of messages can be received and sent on all ports.
debug uldp {hello|probe|echo| unidir|all} [receive|send] interface ethernet <*IFname*> **no debug uldp {hello|probe|echo| unidir|all} [receive|send] interface ethernet <*IFname*>**	Enable or disable the content detail of a particular type of messages can be received and sent on the specified port.

4.3 ULDP Function Typical Examples



Fiber Cross Connection

In the network topology in Graph, port g1/0/1 and port g1/0/2 of switch A as well as port g1/0/3 and port g1/0/4 of switch B are all fiber ports. And the connection is cross connection. The physical layer is connected and works normally, but the data link layer is abnormal. ULDP can discover and disable this kind of error state of link. The final result is that port g1/0/1, g1/0/2 of switch A and port g1/0/3, g1/0/4 of switch B are all shut down by ULDP. Only when the connection is correct, can the ports work normally (won't be shut down).

Switch A configuration sequence:

SwitchA(config)#uldp enable

SwitchA(config)#interface ethernet 1/0/1

SwitchA (Config-If-Ethernet1/0/1)#uldp enable

SwitchA (Config-If-Ethernet1/0/1)#exit

SwitchA(config)#interface ethernet1/0/2

SwitchA(Config-If-Ethernet1/0/2)#uldp enable

Switch B configuration sequence:

SwitchB(config)#uldp enable

SwitchB(config)#interface ethernet1/0/3

SwitchB(Config-If-Ethernet1/0/3)#uldp enable

SwitchB(Config-If-Ethernet1/0/3)#exit

SwitchB(config)#interface ethernet1/0/4

SwitchB(Config-If-Ethernet1/0/4)#uldp enable

As a result, port g1/0/1, g1/0/2 of switch A are all shut down by ULDP, and there is notification information on the CRT terminal of PC1.

%Oct 29 11:09:50 2007 A unidirectional link is detected! Port Ethernet1/0/1 need to be shutted down!

%Oct 29 11:09:50 2007 Unidirectional port Ethernet1/0/1 shut down!
%Oct 29 11:09:50 2007 A unidirectional link is detected! Port Ethernet1/0/2 need to be shutted down!
%Oct 29 11:09:50 2007 Unidirectional port Ethernet1/0/2 shutted down!
Port g1/0/3, and port g1/0/4 of switch B are all shut down by ULDP, and there is notification information on the CRT terminal of PC2.
%Oct 29 11:09:50 2007 A unidirectional link is detected! Port Ethernet1/0/3 need to be shutted down!
%Oct 29 11:09:50 2007 Unidirectional port Ethernet1/0/3 shutted down!
%Oct 29 11:09:50 2007 A unidirectional link is detected! Port Ethernet1/0/4 need to be shutted down!
%Oct 29 11:09:50 2007 Unidirectional port Ethernet1/0/4 shutted down!

4.4 ULDP Troubleshooting

Configuration Notice:
In order to ensure that ULDP can discover that the one of fiber ports has not connected or the ports are incorrectly cross connected, the ports have to work in duplex mode and have the same rate.
If the automatic negotiation mechanism of the fiber ports with one port misconnected decides the working mode and rate of the ports, ULDP won't take effect no matter enabled or not. In such situation, the port is considered as "Down".
In order to make sure that neighbors can be correctly created and unidirectional links can be correctly discovered, it is required that both end of the link should enable ULDP, using the same authentication method and password. At present, no password is needed on both ends.
The hello interval of sending hello messages can be changed (it is10 seconds by default and ranges from 5 to 100 seconds) so that ULDP can respond faster to connection errors of links in different network environments. But this interval should be less than 1/3 of the STP convergence time. If the interval is too long, a STP loop will be generated before ULDP discovers and shuts down the unidirectional connection port. If the interval is too short, the network burden on the port will be increased, which means a reduced bandwidth.
ULDP does not handle any LACP event. It treats every link of TRUNK group (like Port-channal, TRUNK ports) as independent, and handles each of them respectively.
ULDP does not compact with similar protocols of other vendors, which means users can not use ULDP on one end and use other similar protocols on the other end.
ULDP function is disabled by default. After globally enabling ULDP function, the debug switch can be enabled simultaneously to check the debug information. There are several DEBUG commands provided to print debug information, such as information of events, state machine, errors and messages. Different types of message information can also be printed according to different parameters.

The Recovery timer is disabled by default and will only be enabled when the users have configured recovery time (30-86400 seconds).

Reset command and reset mechanism can only reset the ports automatically shut down by ULDP. The ports shut down manually by users or by other modules won't be reset by ULDP.

Chapter 5 LLDP Function Operation Configuration

5.1 Introduction to LLDP Function

Link Layer Discovery Protocol (LLDP) is a new protocol defined in 802.1ab. It enables neighbor devices to send notices of their own state to other devices, and enables all ports of every device to store information about them. If necessary, the ports can also send update information to the neighbor devices directly connected to them, and those neighbor devices will store the information in standard SNMP MIBs. The network management system can check the layer-two connection state from MIB. LLDP won't configure or control network elements or flows, but only report the configuration of layer-two. Another content of 802.1ab is to utilizing the information provided by LLDP to find the conflicts in layer-two. IEEE now uses the existing physical topology, interfaces and Entity MIBs of IETF.

To simplify, LLDP is a neighbor discovery protocol. It defines a standard method for Ethernet devices, such as switches, routers and WLAN access points, to enable them to notify their existence to other nodes in the network and store the discovery information of all neighbor devices. For example, the detail information of the device configuration and discovery can both use this protocol to advertise.

In specific, LLDP defines a general advertisement information set, a transportation advertisement protocol and a method to store the received advertisement information. The device to advertise its own information can put multiple pieces of advertisement information in one LAN data packet to transport. The type of transportation is the type length value (TLV) field. All devices supporting LLDP have to support device ID and port ID advertisement, but it is assumed that, most devices should also support system name, system description and system performance advertisement. System name and system description advertisement can also provide useful information for collecting network flow data. System description advertisement can include data such as the full name of the advertising device, hardware type of system, the version information of software operation system and so on.

802.1AB Link Layer Discovery Protocol will make searching the problems in an enterprise network an easier process and can strengthen the ability of network management tools to discover and maintain accurate network topology structure.

Many kinds of network management software use "Automated Discovery" function to trace the change and condition of topology, but most of them can reach layer-three and classify the devices into all IP subnets at best. This kind of data are very primitive, only referring to basic events like the adding and removing of relative devices instead of details about where and how these devices operate with the network.

Layer 2 discovery covers information like which devices have which ports, which switches

connect to other devices and so on, it can also display the routs between clients, switches, routers, application servers and network servers. Such details will be very meaningful for schedule and investigate the source of network failure.

LLDP will be a very useful management tool, providing accurate information about network mirroring, flow data and searching network problems.

5.2 LLDP Function Configuration Task Sequence

1. Globally enable LLDP function
2. Configure the port-based LLDP function switch
3. Configure the operating state of port LLDP
4. Configure the intervals of LLDP updating messages
5. Configure the aging time multiplier of LLDP messages
6. Configure the sending delay of updating messages
7. Configure the intervals of sending Trap messages
8. Configure to enable the Trap function of the port
9. Configure the optional information-sending attribute of the port
10. Configure the size of space to store Remote Table of the port
11. Configure the type of operation when the Remote Table of the port is full
12. Display and debug the relative information of LLDP

1. Globally enable LLDP function

Command	Explanation
Global Mode	
lldp enable **lldp disable**	Globally enable or disable LLDP function.

2. Configure the port-base LLDP function switch

Command	Explanation
Port Mode	
lldp enable **lldp disable**	Configure the port-base LLDP function switch.

3. Configure the operating state of port LLDP

Command	Explanation
Port Mode	
lldp mode (send|receive|both|disable)	Configure the operating state of port LLDP.

4. Configure the intervals of LLDP updating messages

Command	Explanation
Global Mode	
lldp tx-interval <*integer*> **no lldp tx-interval**	Configure the intervals of LLDP updating messages as the specified value or default value.

5. Configure the aging time multiplier of LLDP messages

Command	Explanation
Global Mode	
lldp msgTxHold <*value*> **no lldp msgTxHold**	Configure the aging time multiplier of LLDP messages as the specified value or default value.

6. Configure the sending delay of updating messages

Command	Explanation
Global Mode	
lldp transmit delay <*seconds*> **no lldp transmit delay**	Configure the sending delay of updating messages as the specified value or default value.

7. Configure the intervals of sending Trap messages

Command	Explanation
Global Mode	
lldp notification interval <*seconds*> **no lldp notification interval**	Configure the intervals of sending Trap messages as the specified value or default value.

8. Configure to enable the Trap function of the port

Command	Explanation
Port Configuration Mode	
lldp trap <enable|disable>	Enable or disable the Trap function of the port.

9. Configure the optional information-sending attribute of the port

Command	Explanation
Port Configuration Mode	
lldp transmit optional tlv [portDesc]	Configure the optional information-sending

[sysName] [sysDesc] [sysCap] **no lldp transmit optional tlv**	attribute of the port as the option value of default values.

10. Configure the size of space to store Remote Table of the port

Command	Explanation
Port Configuration Mode	
lldp neighbors max-num < *value* > **no lldp neighbors max-num**	Configure the size of space to store Remote Table of the port as the specified value or default value.

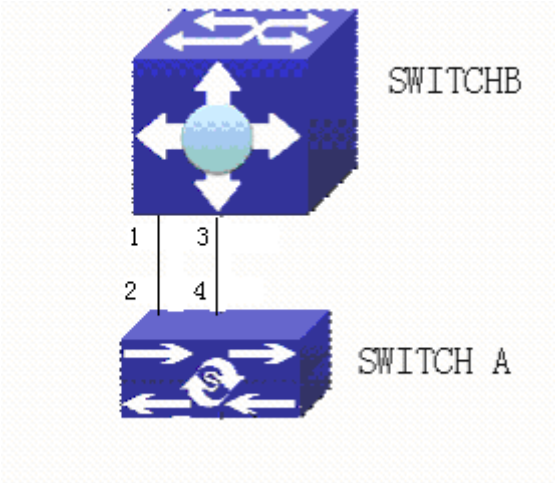
11. Configure the type of operation when the Remote Table of the port is full

Command	Explanation
Port Configuration Mode	
lldp tooManyNeighbors {discard | delete}	Configure the type of operation when the Remote Table of the port is full.

12. Display and debug the relative information of LLDP

Command	Explanation
Admin, Global Mode	
show lldp	Display the current LLDP configuration information.
show lldp interface ethernet <*IFNAME*>	Display the LLDP configuration information of the current port.
show lldp traffic	Display the information of all kinds of counters.
show lldp neighbors interface ethernet < *IFNAME* >	Display the information of LLDP neighbors of the current port.
show debugging lldp	Display all ports with LLDP debug enabled.
Admin Mode	
debug lldp **no debug lldp**	Enable or disable the DEBUG switch.
debug lldp packets interface ethernet <*IFNAME*> **no debug lldp packets interface ethernet <*IFNAME*>**	Enable or disable the DEBUG packet-receiving and sending function in port or global mode.
Port configuration mode	
clear lldp remote-table	Clear Remote-table of the port.

5.3 LLDP Function Typical Example



LLDP Function Typical Configuration Example

In the network topology graph above, the port 1,3 of switch B are connected to port 2,4 of switch A. Port 1 of switch B is configured to message-receiving-only mode, Option TLV of port 4 of switch A is configured as portDes and SysCap.
switch A configuration task sequence:
SwitchA(config)# lldp enable
SwitchA(config)#interface ethernet 1/0/4
SwitchA(Config-If-Ethernet1/0/4)# lldp transmit optional tlv portDesc sysCap
SwitchA(Config-If-Ethernet1/0/4)exit

switch B configuration task sequence:
SwitchB(config)#lldp enable
SwitchB(config)#interface ethernet1/0/1
SwitchB(Config-If-Ethernet1/0/1)# lldp mode receive
SwitchB(Config-If-Ethernet1/0/1)#exit

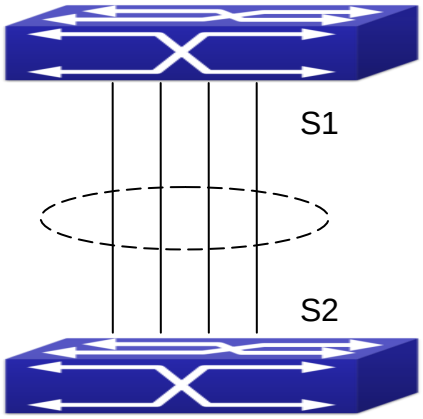
5.4 LLDP Function Troubleshooting

LLDP function is disabled by default. After enabling the global switch of LLDP, users can enable the debug switch “**debug lldp**” simultaneously to check debug information.
Using “show” function of LLDP function can display the configuration information in global or port configuration mode.

Chapter 6 Port Channel Configuration

6.1 Introduction to Port Channel

To understand Port Channel, Port Group should be introduced first. Port Group is a group of physical ports in the configuration level; only physical ports in the Port Group can take part in link aggregation and become a member port of a Port Channel. Logically, Port Group is not a port but a port sequence. Under certain conditions, physical ports in a Port Group perform port aggregation to form a Port Channel that has all the properties of a logical port, therefore it becomes an independent logical port. Port aggregation is a process of logical abstraction to abstract a set of ports (port sequence) with the same properties to a logical port. Port Channel is a collection of physical ports and used logically as one physical port. Port Channel can be used as a normal port by the user, and can not only add network's bandwidth, but also provide link backup. Port aggregation is usually used when the switch is connected to routers, PCs or other switches.



Port aggregation

As shown in the above, S1 is aggregated to a Port Channel, the bandwidth of this Port Channel is the total of all the four ports. If traffic from S1 needs to be transferred to S2 through the Port Channel, traffic allocation calculation will be performed based on the source MAC address and the lowest bit of target MAC address. The calculation result will decide which port to convey the traffic. If a port in Port Channel fails, the other ports will undertake traffic of that port through a traffic allocation algorithm. This algorithm is carried out by the hardware.

Switch offers two methods for configuring port aggregation: manual Port Channel creation and LACP (Link Aggregation Control Protocol) dynamic Port Channel creation. Port aggregation can only be performed on ports in full-duplex mode.

For Port Channel to work properly, member ports of the Port Channel must have the same properties as follows:

All ports are in full-duplex mode.

All Ports are of the same speed.
All ports are Access ports and belong to the same VLAN or are all TRUNK ports, or are all Hybrid ports.
If the ports are all TRUNK ports or Hybrid ports, then their “Allowed VLAN” and “Native VLAN” property should also be the same.
If Port Channel is configured manually or dynamically on switch, the system will automatically set the port with the smallest number to be Master Port of the Port Channel. If the spanning tree function is enabled in the switch, the spanning tree protocol will regard Port Channel as a logical port and send BPDU frames via the master port.
Port aggregation is closely related with switch hardware. Switch allow physical port aggregation of any two switches, maximum 128 groups and 8 ports in each port group are supported.
Once ports are aggregated, they can be used as a normal port. Switch have a built-in aggregation interface configuration mode, the user can perform related configuration in this mode just like in the VLAN and physical interface configuration mode.

6.2 Brief Introduction to LACP

LACP (Link Aggregation Control Protocol) is a kind of protocol based on IEEE802.3ad standard to implement the link dynamic aggregation. LACP protocol uses LACPDU (Link Aggregation Control Protocol Data Unit) to exchange the information with the other end.
After LACP protocol of the port is enabled, this port will send LACPDU to the other end to notify the system priority, the MAC address of the system, the priority of the port, the port ID and the operation Key. After the other end receives the information, the information is compared with the saving information of other ports to select the port which can be aggregated, accordingly, both sides can reach an agreement about the ports join or exit the dynamic aggregation group. The operation Key is created by LACP protocol according to the combination of configuration (speed, duplex, basic configuration, management Key) of the ports to be aggregated.
After the dynamic aggregation port enables LACP protocol, the management Key is 0 by default. After the static aggregation port enables LACP, the management Key of the port is the same with the ID of the aggregation group.
For the dynamic aggregation group, the members of the same group have the same operation Key, for the static aggregation group, the ports of Active have the same operation Key.
The port aggregation is that multi-ports are aggregated to form an aggregation group, so as to implement the out/in load balance in each member port of the aggregation group and provides the better reliability.

6.2.1 Static LACP Aggregation

Static LACP aggregation is enforced by users configuration, and do not enable LACP protocol.

When configuring static LACP aggregation, use “on” mode to force the port to enter the aggregation group.

6.2.2 Dynamic LACP Aggregation

1. The summary of the dynamic LACP aggregation

Dynamic LACP aggregation is an aggregation created/deleted by the system automatically, it does not allow the user to add or delete the member ports of the dynamic LACP aggregation. The ports which have the same attribute of speed and duplex, are connected to the same device, have the same basic configuration, can be dynamically aggregated together. Even if only one port can create the dynamic aggregation, that is the single port aggregation. In the dynamic aggregation, LACP protocol of the port is at the enable state.

2. The port state of the dynamic aggregation group

In dynamic aggregation group, the ports have two states: selected or standby. Both selected ports and standby ports can receive and send LACP protocol, but standby ports can not forward the data packets.

Because the limitation of the max port number in the aggregation group, if the current number of the member ports exceeds the limitation of the max port number, then the system of this end will negotiates with the other end to decide the port state according to the port ID. The negotiation steps are as follows:

Compare ID of the devices (the priority of the system + the MAC address of the system). First, compare the priority of the systems, if they are same, then compare the MAC address of the systems. The end with a small device ID has the high priority.

Compare the ID of the ports (the priority of the port + the ID of the port). For each port in the side of the device which has the high device priority, first, compare the priority of the ports, if the priorities are same, then compare the ID of the ports. The port with a small port ID is selected, and the others become the standby ports.

In an aggregation group, the port which has the smallest port ID and is at the selected state will be the master port, the other ports at the selected state will be the member port.

6.3 Port Channel Configuration Task List

1. Create a port group in Global Mode
2. Add ports to the specified group from the Port Mode of respective ports
3. Enter port-channel configuration mode
4. Set load-balance method for Port-group
5. Set the system priority of LACP protocol
6. Set the port priority of the current port in LACP protocol
7. Set the timeout mode of the current port in LACP protocol

1. Creating a port group

Command	Explanation
Global Mode	
port-group <port-group-number> **no port-group <port-group-number>**	Create or delete a port group.

2. Add physical ports to the port group

Command	Explanation
Port Mode	
port-group <*port-group-number*> mode {active | passive | on} **no port-group**	Add the ports to the port group and set their mode.

3. Enter port-channel configuration mode.

Command	Explanation
Global Mode	
interface port-channel <*port-channel-number*>	Enter port-channel configuration mode.

4. Set load-balance method for switch

Command	Explanation
Global configuration mode	
load-balance {dst-src-mac | dst-src-ip | dst-src-mac-ip}	Set load-balance for switch, it takes effect on port-group and ECMP function at the same time.

5. Set the system priority of LACP protocol

Command	Explanation
Global mode	
lacp system-priority <*system-priority*> **no lacp system-priority**	Set the system priority of LACP protocol, the no command restores the default value.

6. Set the port priority of the current port in LACP protocol

Command	Explanation
Port mode	
lacp port-priority <*port-priority*>	Set the port priority in LACP protocol.

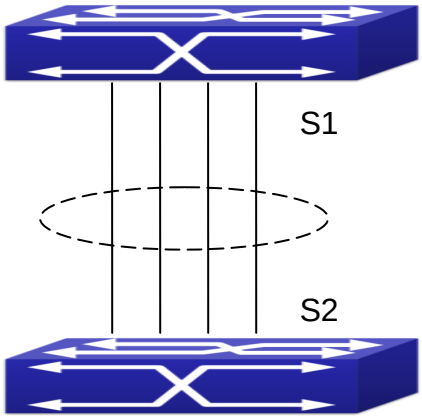
no lacp port-priority	The no command restores the default value.

7. Set the timeout mode of the current port in LACP protocol

Command	Explanation
Port mode	
lacp timeout {short | long} **no lacp timeout**	Set the timeout mode in LACP protocol. The no command restores the default value.

6.4 Port Channel Examples

Scenario 1: Configuring Port Channel in LACP.



Configure Port Channel in LACP

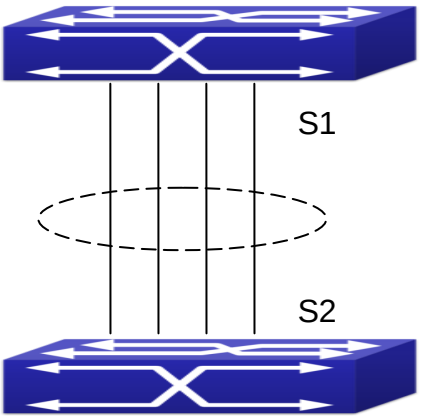
The switches in the description below are all switch and as shown in the figure, ports 1, 2, 3, 4 of S1 are access ports and add them to group1 with active mode. Ports 6, 8, 9, 10 of S2 are access ports and add them to group2 with passive mode. All the ports should be connected with cables.

The configuration steps are listed below:
Switch1#config
Switch1(config)#interface ethernet 1/0/1-4
Switch1(Config-If-Port-Range)#port-group 1 mode active
Switch1(Config-If-Port-Range)#exit
Switch1(config)#interface port-channel 1
Switch1(Config-If-Port-Channel1)#

Switch2#config
Switch2(config)#port-group 2
Switch2(config)#interface ethernet 1/0/6
Switch2(Config-If-Ethernet1/0/6)#port-group 2 mode passive
Switch2(Config-If-Ethernet1/0/6)#exit
Switch2(config)#interface ethernet 1/0/8-10
Switch2(Config-If-Port-Range)#port-group 2 mode passive
Switch2(Config-If-Port-Range)#exit
Switch2(config)#interface port-channel 2
Switch2(Config-If-Port-Channel2)#

Configuration result:
Shell prompts ports aggregated successfully after a while, now ports 1, 2, 3, 4 of S1 form an aggregated port named “Port-Channel1”, ports 6, 8, 9, 10 of S2 form an aggregated port named “Port-Channel2”; can be configured in their respective aggregated port mode.

Scenario 2: Configuring Port Channel in ON mode.



Configure Port Channel in ON mode

As shown in the figure, ports 1, 2, 3, 4 of S1 are access ports and add them to group1 with “on” mode. Ports 6, 8, 9, 10 of S2 are access ports and add them to group2 with “on” mode.

The configuration steps are listed below:
Switch1#config
Switch1(config)#interface ethernet 1/0/1
Switch1(Config-If-Ethernet1/0/1)#port-group 1 mode on
Switch1(Config-If-Ethernet1/0/1)#exit
Switch1(config)#interface ethernet 1/0/2
Switch1(Config-If-Ethernet1/0/2)#port-group 1 mode on
Switch1(Config-If-Ethernet1/0/2)#exit

```
Switch1(config)#interface ethernet 1/0/3
Switch1(Config-If-Ethernet1/0/3)#port-group 1 mode on
Switch1(Config-If-Ethernet1/0/3)#exit
Switch1(config)#interface ethernet 1/0/4
Switch1(Config-If-Ethernet1/0/4)#port-group 1 mode on
Switch1(Config-If-Ethernet1/0/4)#exit

Switch2#config
Switch2(config)#port-group 2
Switch2(config)#interface ethernet 1/0/6
Switch2(Config-If-Ethernet1/0/6)#port-group 2 mode on
Switch2(Config-If-Ethernet1/0/6)#exit
Switch2(config)#interface ethernet 1/0/8-10
Switch2(Config-If-Port-Range)#port-group 2 mode on
Switch2(Config-If-Port-Range)#exit
```

Configuration result:

Add ports 1, 2, 3, 4 of S1 to port-group1 in order, and we can see a group in “on” mode is completely joined forcedly, switch in other ends won’t exchange LACP PDU to complete aggregation. Aggregation finishes immediately when the command to add port 2 to port-group 1 is entered, port 1 and port 2 aggregate to be port-channel 1, when port 3 joins port-group 1, port-channel 1 of port 1 and 2 are ungrouped and re-aggregate with port 3 to form port-channel 1, when port 4 joins port-group 1, port-channel 1 of port 1, 2 and 3 are ungrouped and re-aggregate with port 4 to form port-channel 1. (It should be noted that whenever a new port joins in an aggregated port group, the group will be ungrouped first and re-aggregated to form a new group.) Now all four ports in both S1 and S2 are aggregated in “on” mode and become an aggregated port respectively.

6.5 Port Channel Troubleshooting

If problems occur when configuring port aggregation, please first check the following for causes.

Ensure all ports in a port group have the same properties, i.e., whether they are in full-duplex mode, forced to the same speed, and have the same VLAN properties, etc. If inconsistency occurs, make corrections.

Some commands cannot be used on a port in port-channel, such as arp, bandwidth, ip, ip-forward, etc.

Chapter 7 Jumbo Configuration

7.1 Introduction to Jumbo

So far the Jumbo (Jumbo Frame) has not reach a determined standard in the industry (including the format and length of the frame). Normally frames sized within 1519-9000 should be considered jumbo frame. Networks with jumbo frames will increase the speed of the whole network by 2% to 5%. Technically the Jumbo is just a lengthened frame sent and received by the switch. However considering the length of Jumbo frames, they will not be sent to CPU. We discard the Jumbo frames sent to CPU in the packet receiving process.

7.2 Jumbo Configuration Task Sequence

1. Configure enable Jumbo function

Command	Explanation
Global Mode	
jumbo enable [<mtu-value>] **no jumbo enable**	Enable the receiving/sending function of JUMBO frame. The no command disables sending and receiving function of JUMBO frames.

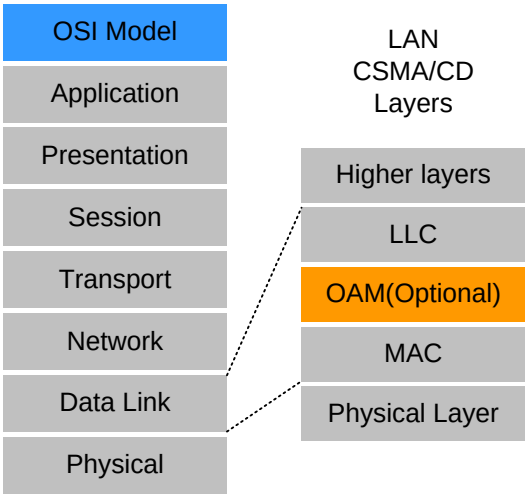
Chapter 8 EFM OAM Configuration

8.1 Introduction to EFM OAM

Ethernet is designed for Local Area Network at the beginning, but link length and network scope is extended rapidly while Ethernet is also applied to Metropolitan Area Network and Wide Area Network along with development. Due to lack the effectively management mechanism, it affects Ethernet application to Metropolitan Area Network and Wide Area Network, implementing OAM on Ethernet becomes a necessary development trend.

There are four protocol standards about Ethernet OAM, they are 802.3ah (EFM OAM), 802.3ag (CFM), E-LMI and Y.1731. EFM OAM and CFM are set for IEEE organization. EFM OAM works in data link layer to validly discover and manage the data link status of rock-bottom. Using EFM OAM can effectively advance management and maintenance for Ethernet to ensure the stable network operation. CFM is used for monitoring the whole network connectivity and locating the fault in access aggregation network layer. Compare with CFM, Y.1731 standard set by ITU (International Telecommunications Union) is more powerful. E-LMI standard set by MEF is only applied to UNI. So above protocols can be used to different network topology and management, between them exist the complementary relation.

EFM OAM (Ethernet in the First Mile Operation, Administration and Maintenance) works in data link layer of OSI model to implement the relative functions through OAM sublayer, figure is as bleow:



OAM location in OSI model

OAM protocol data units (OAMPDU) use destination MAC address 01-80-c2-00-00-02 of protocol, the max transmission rate is 10Pkt/s.

EFM OAM is established on the basis of OAM connection, it provides a link operation management mechanism such as link monitoring, remote fault detection and remote loopback testing, the simple introduction for EFM OAM in the following:

1. Ethernet OAM connection establishment

Ethernet OAM entity discovers remote OAM entities and establishes sessions with them by exchanging Information OAMPDUs. EFM OAM can operate in two modes: active mode and passive mode. One session can only be established by the OAM entity working in the active mode and ones working in the passive mode need to wait until it receives the connection request. After an Ethernet OAM connection is established, the Ethernet OAM entities on both sides exchange Information OAMPDUs continuously to keep the valid Ethernet OAM connection. If an Ethernet OAM entity receives no Information OAMPDU for five seconds, the Ethernet OAM connection is disconnected.

2. Link Monitoring

Fault detection in an Ethernet is difficult, especially when the physical connection in the network is not disconnected but network performance is degrading gradually. Link monitoring is used to detect and discover link faults in various environments. EFM OAM implements link monitoring through the exchange of Event Notification OAMPDUs. When detecting a link error event, the local OAM entity sends an Event Notification OAMPDU to notify the remote OAM entity. At the same time it will log information and send SNMP Trap to the network management system. While OAM entity on the other side receives the notification, it will also log and report it. With the log information, network administrators can keep track of network status in time.

The link event monitored by EFM OAM means that the link happens the error event, including Errored symbol period event, Errored frame event, Errored frame period event, Errored frame seconds event.

Errored symbol period event: The errored symbol number can not be less than the low threshold. (Symbol: the min data transmission unit of physical medium. It is unique for coding system, the symbols may be different for different physical mediums, symbol rate means the changed time of electron status per second.)

Errored frame period event: Specifying N is frame period, the errored frame number within the period of receiving N frames can not be less than the low threshold. (Errored frame: Receiving the errored frame detected by CRC.)

Errored frame event: The number of detected error frames over M seconds can not be less than the low threshold.

Errored frame seconds event: The number of error frame seconds detected over M seconds can not be less than the low threshold. (Errored frame second: Receiving an errored frame at least in a second.)

3. Remote Fault Detection

In a network where traffic is interrupted due to device failures or unavailability, the flag field defined in Ethernet OAMPDUs allows an Ethernet OAM entity to send fault information to its

peer. As Information OAMPDUs are exchanged continuously across established OAM connections, an Ethernet OAM entity can inform one of its OAM peers of link faults through Information OAMPDUs. Therefore, the network administrator can keep track of link status in time through the log information and troubleshoot in time.

There are three kinds of link faults for Information OAMPDU, they are Critical Event, Dying Gasp and Link Fault, and their definitions are different for each manufacturer, here the definitions are as below:

Critical Event: EFM OAM function of port is disabled.

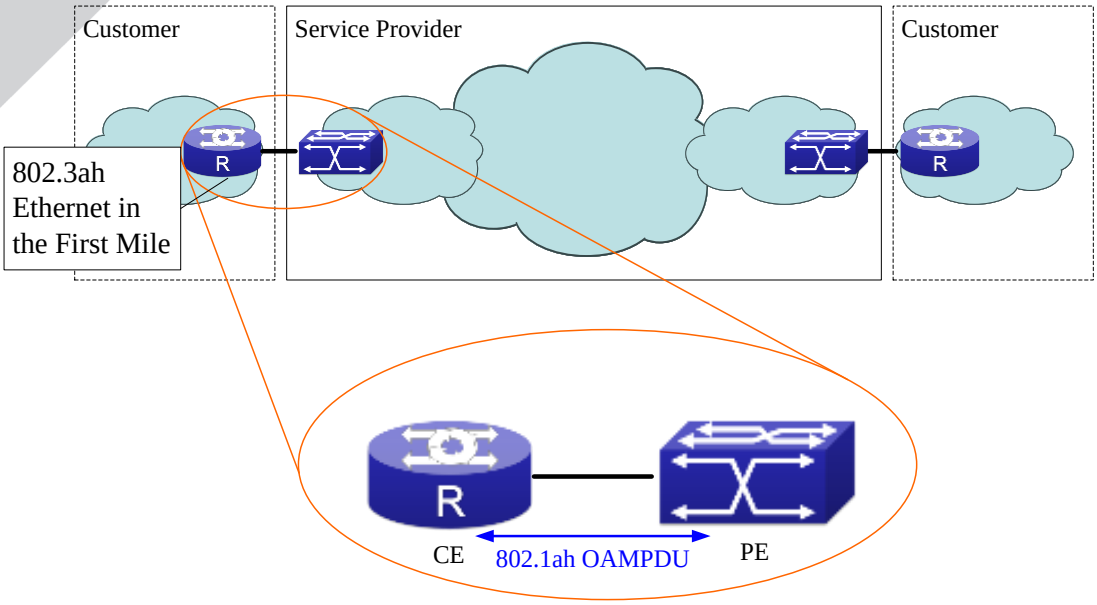
Link Fault: The number of unidirectional operation or fault can not be less than the high threshold in local. Unidirectional Operation means unidirectional link can not work normally on full-duplex link without autonegotiaction. EFM OAM can detect the fault and inform the remote OAM peers through sending Information OAMPDU.

Dying Gasp: There is no definition present. Although device does not generate Dying Gasp OAMPDU, it still receives and processes such OAMPDU sent by its peer.

4. Remote loopback testing

Remote loopback testing is available only after an Ethernet OAM connection is established. With remote loopback enabled, operating Ethernet OAM entity in active mode issues remote loopback requests and the peer responds to them. If the peer operates in loopback mode, it returns all packets except Ethernet OAMPDUs to the senders along the original paths. Performing remote loopback testing periodically helps to detect network faults in time. Furthermore, performing remote loopback testing by network segments helps to locate network faults. Note: The communication will not be processed normally in remote loopback mode.

Typical EFM OAM application topology is in the following, it is used for point-to-point link and emulational IEEE 802.3 point-to-point link. Device enables EFM OAM through point-to-point connection to monitor the link fault in the First Mile with Ethernet access. For user, the connection between user to telecommunication is “the First Mile”, for service provider, it is “the Last Mile”.



Typical OAM application topology

8.2 EFM OAM Configuration

EFM OAM configuration task list

1. Enable EFM OAM function of port
2. Configure link monitor
3. Configure remote failure
4. Enable EFM OAM loopback of port

Note: it needs to enable OAM first when configuring OAM parameters.

1. Enable EFM OAM function of port

Command	Explanation
Port mode	
ethernet-oam mode {active | passive}	Configure work mode of EFM OAM, default is active mode.
ethernet-oam **no ethernet-oam**	Enable EFM OAM of port, no command disables EFM OAM of port.
ethernet-oam period <seconds> **no ethernet-oam period**	Configure transmission period of OAMPDU (optional), no command restores the default value.
ethernet-oam timeout <seconds> **no ethernet-oam timeout**	Configure timeout of EFM OAM connection, no command restores the default value.

2. Configure link monitor

Command	Explanation
Port mode	
ethernet-oam link-monitor **no ethernet-oam link-monitor**	Enable link monitor of EFM OAM, no command disables link monitor.
ethernet-oam errored-symbol-period {threshold low <low-symbols> | window <seconds>} **no ethernet-oam errored-symbol-period {threshold low | window }**	Configure the low threshold and window period of errored symbol period event, no command resotores the default value. (optional)
ethernet-oam errored-frame-period {threshold low <low-frames> | window <seconds>} **no ethernet-oam errored-frame-period {threshold low | window }**	Configure the low threshold and window period of errored frame period event, no command resotores the default value.
ethernet-oam errored-frame {threshold low <low-frames> | window <seconds>} **no ethernet-oam errored-frame {threshold low | window }**	Configure the low threshold and window period of errored frame event, no command resotores the default value. (optional)
ethernet-oam errored-frame-seconds {threshold low <low-frame-seconds> | window <seconds>} **no ethernet-oam errored-frame-seconds {threshold low | window }**	Configure the low threshold and window period of errored frame seconds event, no command resotores the default value. (optional)

3. Configure remote failure

Command	Explanation
Port mode	
ethernet-oam remote-failure **no ethernet-oam remote-failure**	Enable remote failure detection of EFM OAM (failure means critical-event or link-fault event of the local), no command disables the function. (optional)
ethernet-oam errored-symbol-period threshold high {high-symbols | none} **no ethernet-oam errored-symbol-period threshold high**	Configure the high threshold of errored symbol period event, no command restores the default value. (optional)
ethernet-oam errored-frame-period threshold high {high-frames | none} **no ethernet-oam errored-frame-period threshold high**	Configure the high threshold of errored frame period event, no command restores the default value. (optional)

ethernet-oam errored-frame threshold high {high-frames | none} **no ethernet-oam errored-frame threshold high**	Configure the high threshold of errored frame event, no command restores the default value. (optional)
ethernet-oam errored-frame-seconds threshold high {high-frame-seconds | none} **no ethernet-oam errored-frame-seconds threshold high**	Configure the high threshold of errored frame seconds event, no command restores the default value. (optional)

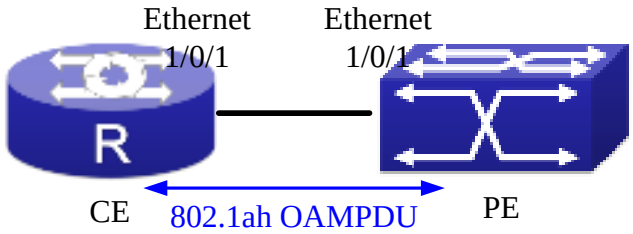
4. Enable EFM OAM loopback of port

Command	Explanation
Port mode	
ethernet-oam remote-loopback **no ethernet-oam remote-loopback**	Enable remote EFM OAM entity to enter OAM loopback mode (its peer needs to configure OAM loopback supporting), no command cancels remote OAM loopback.
ethernet-oam remote-loopback supported **no ethernet-oam remote-loopback supported**	Enable remote loopback supporting of port, no command cancels remote loopback supporting of port.

8.3 EFM OAM Example

Example:

CE and PE devices with point-to-point link enable EFM OAM to monitor “the First Mile” link performance. It will report the log information to network management system when occurring fault event and use remote loopback function to detect the link in necessary instance



Typical OAM application topology

Configuration procedure: (Omitting SNMP and Log configuration in the following)

Configuration on CE:

CE(config)#interface ethernet 1/0/1

CE (config-if-ethernet1/0/1)#ethernet-oam mode passive

CE (config-if-ethernet1/0/1)#ethernet-oam

CE (config-if-ethernet1/0/1)#ethernet-oam remote-loopback supported
Other parameters use the default configuration.

Configuration on PE:
PE(config)#interface ethernet 1/0/1
PE (config-if-ethernet1/0/1)#ethernet-oam
Other parameters use the default configuration.

Execute the following command when using remote loopback.
PE(config-if-ethernet1/0/1)#ethernet-oam remote-loopback
Execute the following command to make one of OAM peers exiting OAM loopback after complete detection.
PE(config-if-ethernet1/0/1)# no ethernet-oam remote-loopback
Execute the following command without supporting remote loopback.
CE(config-if-ethernet1/0/1)#no ethernet-oam remote-loopback supported

8.4 EFM OAM Troubleshooting

When using EFM OAM, it occurs the problem, please check whether the problem is resulted by the following reasons:
Check whether OAM entities of two peers of link in passive mode. If so, EFM OAM connection can not be established between two OAM entities.
Ensuring SNMP configuration is correct, or else errored event can not be reported to network management system.
Link does not normally communicate in OAM loopback mode, it should cancel remote loopback in time after detect the link performance.
Ensuring the used board supports remote loopback function.
Port should not configure STP, MRPP, ULPP, Flow Control, loopback detection functions after it enables OAM loopback function, because OAM remote loopback function and these functions are mutually exclusive.
When enabling OAM, the negotiation of the port will be disabled automatically. So the negotiation in the peer of the link must be disabled, otherwise the link connection will unsuccessful. When disabling OAM, the negotiation of the port will be restored. Therefore, to ensure the link connection is normal, the negotiations must be accordant in two peers of the link.
After enabling OAM, when the link negotiations in two peers are successful, the state is up. After the fiber in RX redirection of the peer is pulled out, TX of the peer and RX with OAM are normal, so the port with OAM will be at up state all along.

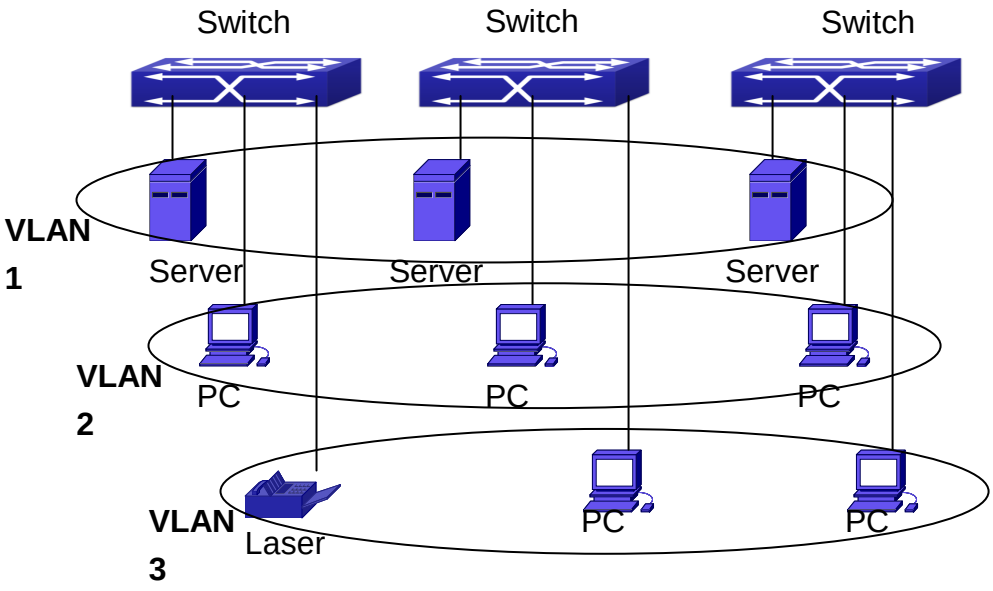
Chapter 9 VLAN Configuration

9.1 VLAN Configuration

9.1.1 Introduction to VLAN

VLAN (Virtual Local Area Network) is a technology that divides the logical addresses of devices within the network to separate network segments basing on functions, applications or management requirements. By this way, virtual workgroups can be formed regardless of the physical location of the devices. IEEE announced IEEE 802.1Q protocol to direct the standardized VLAN implementation, and the VLAN function of switch is implemented following IEEE 802.1Q.

The key idea of VLAN technology is that a large LAN can be partitioned into many separate broadcast domains dynamically to meet the demands.



A VLAN network defined logically

Each broadcast domain is a VLAN. VLANs have the same properties as the physical LANs, except VLAN is a logical partition rather than physical one. Therefore, the partition of VLANs can be performed regardless of physical locations, and the broadcast, multicast and unicast traffic within a VLAN is separated from the other VLANs.

With the aforementioned features, VLAN technology provides us with the following convenience:

Improving network performance

Saving network resources

Simplifying network management

Lowering network cost

Enhancing network security

Switch Ethernet Ports can works in three kinds of modes: Access, Hybrid and Trunk, each mode has a different processing method in forwarding the packets with tagged or untagged.

The ports of Access type only belongs to one VLAN, usually they are used to connect the ports of the computer.

The ports of Trunk type allow multi-VLANs to pass, can receive and send the packets of multi-VLANs. Usually they are used to connect between the switches.

The ports of Hybrid type allow multi-VLANs to pass, can receive and send the packets of multi-VLANs. They can be used to connect between the switches, or to a computer of the user.

Hybrid ports and Trunk ports receive the data with the same process method, but send the data with different method: Hybrid ports can send the packets of multi-VLANs without the VLAN tag, while Trunk ports send the packets of multi-VLANs with the VLAN tag except the port native VLAN.

The switch implements VLAN and GVRP (GARP VLAN Registration Protocol) which are defined by 802.1Q. The chapter will explain the use and the configuration of VLAN and GVRP in detail.

9.1.2 VLAN Configuration Task List

1. Create or delete VLAN
2. Set or delete VLAN name
3. Assign Switch ports for VLAN
4. Set the switch port type
5. Set Trunk port
6. Set Access port
7. Set Hybrid port
8. Enable/Disable VLAN ingress rules on ports
9. Configure Private VLAN
10. Set Private VLAN association
11. Specify internal VLAN ID

1. Create or delete VLAN

Command	Explanation
Global Mode	
vlan WORD **no vlan WORD**	Create/delete VLAN or enter VLAN Mode

2. Set or delete VLAN name

Command	Explanation

VLAN Mode	
name <vlan-name> **no name**	Set or delete VLAN name.

3. Assigning Switch ports for VLAN

Command	Explanation
VLAN Mode	
switchport interface <*interface-list*> **no switchport interface <*interface-list*>**	Assign Switch ports to VLAN.

4. Set the Switch Port Type

Command	Explanation
Port Mode	
switchport mode {trunk | access | hybrid}	Set the current port as Trunk, Access or Hybrid port.

5. Set Trunk port

Command	Explanation
Port Mode	
switchport trunk allowed vlan {WORD | all | add WORD | except WORD | remove WORD} **no switchport trunk allowed vlan**	Set/delete VLAN allowed to be crossed by Trunk. The “no” command restores the default setting.
switchport trunk native vlan <*vlan-id*> **no switchport trunk native vlan**	Set/delete PVID for Trunk port.

6. Set Access port

Command	Explanation
Port Mode	
switchport access vlan <*vlan-id*> **no switchport access vlan**	Add the current port to the specified VLAN. The “no” command restores the default setting.

7. Set Hybrid port

Command	Explanation
Port Mode	
switchport hybrid allowed vlan {WORD | all |	Set/delete the VLAN which is allowed by

add WORD | except WORD | remove WORD} {tag | untag} **no switchport hybrid allowed vlan**	Hybrid port with tag or untag mode.
switchport hybrid native vlan <*vlan-id*> **no switchport hybrid native vlan**	Set/delete PVID of the port.

8. Disable/Enable VLAN Ingress Rules

Command	Explanation
Port Mode	
vlan ingress enable **no vlan ingress enable**	Enable/Disable VLAN ingress rules.

9. Configure Private VLAN

Command	Explanation
VLAN mode	
private-vlan {primary | isolated | community} **no private-vlan**	Configure current VLAN to Private VLAN. The no command deletes private VLAN.

10. Set Private VLAN association

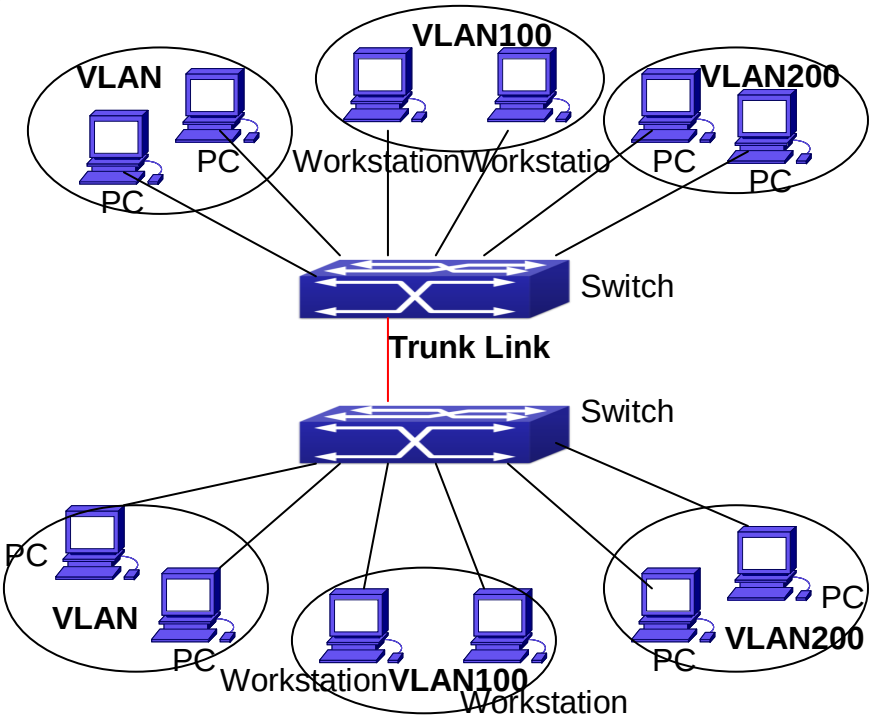
Command	Explanation
VLAN mode	
private-vlan association <secondary-vlan-list> **no private-vlan association**	Set/delete Private VLAN association.

11. Specify internal VLAN ID

Command	Explanation
Global mode	
vlan <2-4094> internal	Specify internal VLAN ID.

9.1.3 Typical VLAN Application

Scenario:



Typical VLAN Application Topology

The existing LAN is required to be partitioned to 3 VLANs due to security and application requirements. The three VLANs are VLAN2, VLAN100 and VLAN200. Those three VLANs are cross two different location A and B. One switch is placed in each site, and cross-location requirement can be met if VLAN traffic can be transferred between the two switches.

Configuration Item	Configuration description
VLAN2	Site A and site B switch port 2 -4.
VLAN100	Site A and site B switch port 5 -7.
VLAN200	Site A and site B switch port 8 -10.
Trunk port	Site A and site B switch port 11.

Connect the Trunk ports of both switches for a Trunk link to convey the cross-switch VLAN traffic; connect all network devices to the other ports of corresponding VLANs.

In this example, port 1 and port 12 is spared and can be used for management port or for other purposes.

The configuration steps are listed below:

Switch A:

Switch(config)#vlan 2

Switch(Config-Vlan2)#switchport interface ethernet 1/0/2-4

Switch(Config-Vlan2)#exit

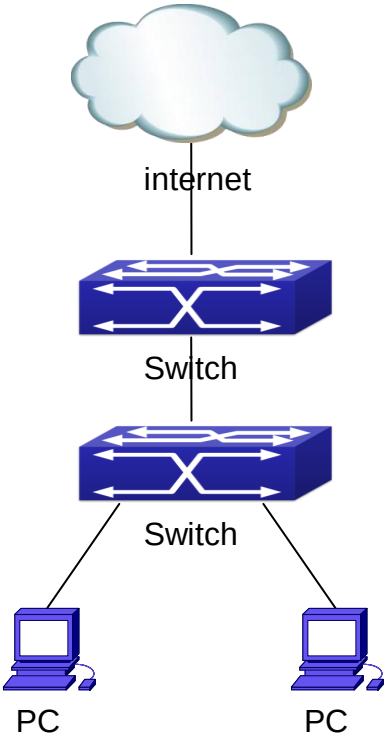
Switch(config)#vlan 100

Switch(Config-Vlan100)#switchport interface ethernet 1/0/5-7

Switch(Config-Vlan100)#exit
Switch(config)#vlan 200
Switch(Config-Vlan200)#switchport interface ethernet 1/0/8-10
Switch(Config-Vlan200)#exit
Switch(config)#interface ethernet 1/0/11
Switch(Config-If-Ethernet1/0/11)#switchport mode trunk
Switch(Config-If-Ethernet1/0/11)#exit
Switch(config)#
Switch B:
Switch(config)#vlan 2
Switch(Config-Vlan2)#switchport interface ethernet 1/0/2-4
Switch(Config-Vlan2)#exit
Switch(config)#vlan 100
Switch(Config-Vlan100)#switchport interface ethernet 1/0/5-7
Switch(Config-Vlan100)#exit
Switch(config)#vlan 200
Switch(Config-Vlan200)#switchport interface ethernet 1/0/8-10
Switch(Config-Vlan200)#exit
Switch(config)#interface ethernet 1/0/11
Switch(Config-If-Ethernet1/0/11)#switchport mode trunk
Switch(Config-If-Ethernet1/0/11)#exit

9.1.4 Typical Application of Hybrid Port

Scenario:



Typical Application of Hybrid Port

PC1 connects to the interface Ethernet 1/0/7 of SwitchB, PC2 connects to the interface Ethernet 1/0/9 of SwitchB, Ethernet 1/0/10 of SwitchA connect to Ethernet 1/0/10 of SwitchB. It is required that PC1 and PC2 can not mutually access due to reason of the security, but PC1 and PC2 can access other network resources through the gateway SwitchA. We can implement this status through Hybrid port.

Configuration items are as follows:

Port	Type	PVID	the VLANs are allowed to pass
Port 1/0/10 of Switch A	Access	10	Allow the packets of VLAN 10 to pass with untag method.
Port 1/0/10 of Switch B	Hybrid	10	Allow the packets of VLAN 7, 9, 10 to pass with untag method.
Port 1/0/7 of Switch B	Hybrid	7	Allow the packets of VLAN 7, 10 to pass with untag method.
Port 1/0/9 of Switch B	Hybrid	9	Allow the packets of VLAN 9, 10 to pass with untag method.

The configuration steps are listed below:

Switch A:

Switch(config)#vlan 10

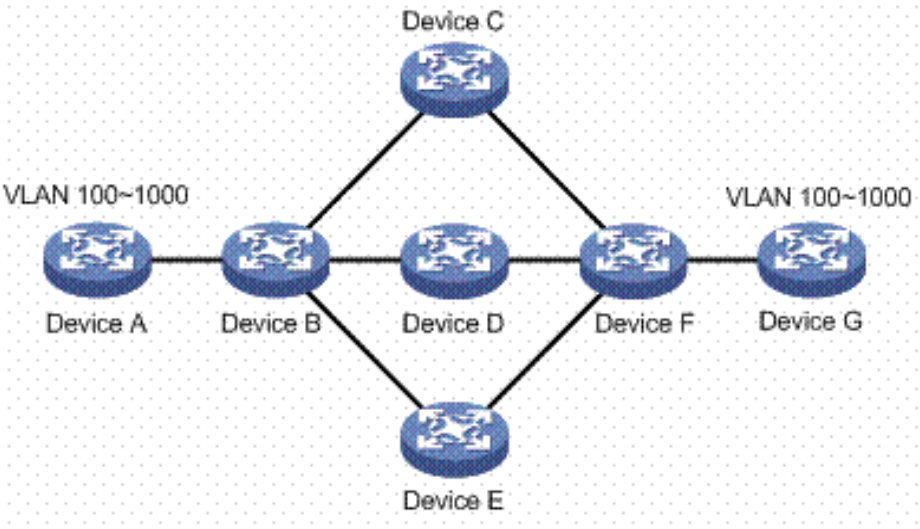
Switch(Config-Vlan10)#switchport interface ethernet 1/0/10

Switch B:
Switch(config)#vlan 7;9;10
Switch(config)#interface ethernet 1/0/7
Switch(Config-If-Ethernet1/0/7)#switchport mode hybrid
Switch(Config-If-Ethernet1/0/7)#switchport hybrid native vlan 7
Switch(Config-If-Ethernet1/0/7)#switchport hybrid allowed vlan 7;10 untag
Switch(Config-If-Ethernet1/0/7)#exit
Switch(Config)#interface Ethernet 1/0/9
Switch(Config-If-Ethernet1/0/9)#switchport mode hybrid
Switch(Config-If-Ethernet1/0/9)#switchport hybrid native vlan 9
Switch(Config-If-Ethernet1/0/9)#switchport hybrid allowed vlan 9;10 untag
Switch(Config-If-Ethernet1/0/9)#exit
Switch(Config)#interface Ethernet 1/0/10
Switch(Config-If-Ethernet1/0/10)#switchport mode hybrid
Switch(Config-If-Ethernet1/0/10)#switchport hybrid native vlan 10
Switch(Config-If-Ethernet1/0/10)#switchport hybrid allowed vlan 7;9;10 untag
Switch(Config-If-Ethernet1/0/10)#exit

9.2 GVRP Configuration

9.2.1 Introduction to GVRP

GVRP, i.e. GARP VLAN Registration Protocol, is an application of GARP (Generic Attribute Registration Protocol). GARP is mainly used to establish an attribute transmission mechanism to transmit attributes, so as to ensure protocol entities registering and deregistering the attribute. According to different transmission attributes, GARP can be divided to many application protocols, such as GMRP and GVRP. Therefore, GVRP is a protocol which transmits VLAN attributes to the whole layer 2 network through GARP protocol.



A typical application scene

A and G switches are not directly connected in layer 2 network; BCDEF are intermediate switches connecting A and G. Switch A and G configure VLAN100-1000 manually while BCDEF switches do not. When GVRP is not enabled, A and G cannot communicate with each other, because intermediate switches without relevant VLANs. However, after GVRP is enabled on all switches, its VLAN attribute transmission mechanism enables the intermediate switches registering the VLANs dynamically, and the VLAN in VLAN100-1000 of A and G can communicate with each other. The VLANs dynamically registered by intermediate switches will be deregistered when deregistering VLAN100-1000 of A and G switches manually. So the same VLAN of two unadjacent switches can communicate mutually through GVRP protocol instead of configuring each intermediate switch manually for achieving the purpose of simplifying VLAN configuration.

9.2.2 GVRP Configuration Task List

GVRP configuration task list:

1. Configure GVRP timer
2. Configure port type
3. Enable GVRP function

1. Configure GVRP timer

Command	Explanation
Global mode	
garp timer join <200-500> **garp timer leave <500-1200>** **garp timer leaveall <5000-60000>** **no garp timer (join | leave | leaveAll)**	Configure leaveall, join and leave timer for GVRP.

2. Configure port type

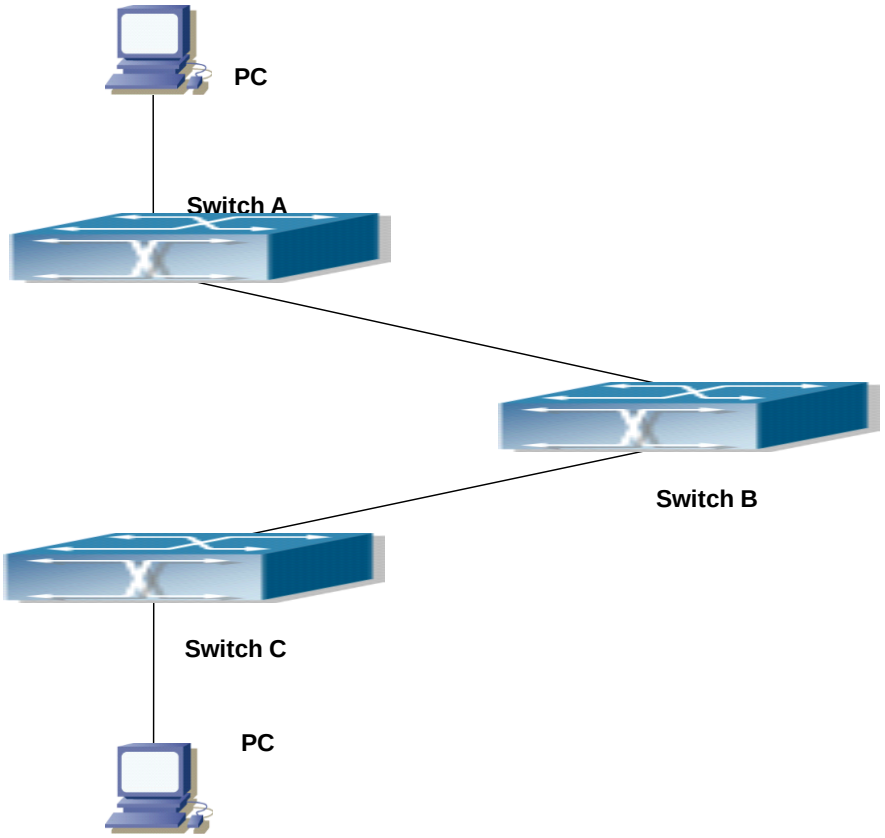
Command	Explanation
Port mode	
gvrp **no gvrp**	Enable/ disable GVRP function of port.

3. Enable GVRP function

Command	Explanation
Global mode	
gvrp **no gvrp**	Enable/ disable the global GVRP function of port.

9.2.3 Example of GVRP

GVRP application:



Typical GVRP Application Topology

To enable dynamic VLAN information register and update among switches, GVRP protocol is to be configured in the switch. Configure GVRP in Switch A, B and C, enable Switch B to learn VLAN100 dynamically so that two workstations connected to VLAN100 in Switch A and C can

communicate with each other through Switch B without static VLAN100 entries.

Configuration Item	Configuration description
VLAN100	Port 2-6 of Switch A and C.
Trunk port	Port 11 of Switch A and C, Port 10, 11 of Switch B.
Global GVRP	Switch A, B, C.
Port GVRP	Port 11 of Switch A and C, Port 10, 11 of Switch B.

Connect two workstations to the VLAN100 ports in switch A and B, connect port 11 of Switch A to port 10 of Switch B, and port 11 of Switch B to port 11 of Switch C.

The configuration steps are listed below:
Switch A:
Switch(config)# gvrp
Switch(config)#vlan 100
Switch(Config-Vlan100)#switchport interface ethernet 1/0/2-6
Switch(Config-Vlan100)#exit
Switch(config)#interface ethernet 1/0/11
Switch(Config-If-Ethernet1/0/11)#switchport mode trunk
Switch(Config-If-Ethernet1/0/11)# gvrp
Switch(Config-If-Ethernet1/0/11)#exit

Switch B:
Switch(config)#gvrp
Switch(config)#interface ethernet 1/0/10
Switch(Config-If-Ethernet1/0/10)#switchport mode trunk
Switch(Config-If-Ethernet1/0/10)# gvrp
Switch(Config-If-Ethernet1/0/10)#exit
Switch(config)#interface ethernet 1/0/11
Switch(Config-If-Ethernet1/0/11)#switchport mode trunk
Switch(Config-If-Ethernet1/0/11)# gvrp
Switch(Config-If-Ethernet1/0/11)#exit

Switch C:
Switch(config)# gvrp
Switch(config)#vlan 100
Switch(Config-Vlan100)#switchport interface ethernet 1/0/2-6
Switch(Config-Vlan100)#exit
Switch(config)#interface ethernet 1/0/11
Switch(Config-If-Ethernet1/0/11)#switchport mode trunk

Switch(Config-If-Ethernet1/0/11)# gvrp
Switch(Config-If-Ethernet1/0/11)#exit

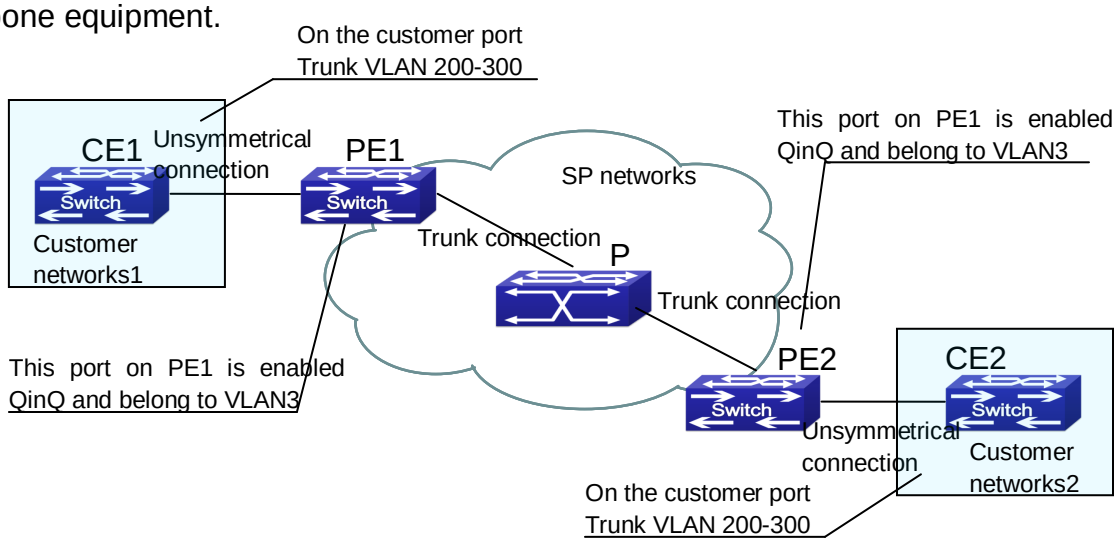
9.2.4 GVRP Troubleshooting

The GARP counter setting for Trunk ports in both ends of Trunk link must be the same, otherwise GVRP will not work normally. It is recommended to avoid enabling GVRP and RSTP at the same time in switch. If GVRP needs to be enabled, RSTP function for the ports must be disabled first.

9.3 Dot1q-tunnel Configuration

9.3.1 Introduction to Dot1q-tunnel

Dot1q-tunnel is also called QinQ (802.1Q-in-802.1Q), which is an expansion of 802.1Q. Its dominating idea is encapsulating the customer VLAN tag (CVLAN tag) to the service provider VLAN tag (SPVLAN tag). Carrying the two VLAN tags the packet is transmitted through the backbone network of the ISP internet, so to provide a simple layer-2 tunnel for the users. It is simple and easy to manage, applicable only by static configuration, and especially adaptive to small office network or small scale metropolitan area network using layer-3 switch as backbone equipment.



Dot1q-tunnel based Internetworking mode

As shown in above, after being enabled on the user port, dot1q-tunnel assigns each user an SPVLAN identification (SPVID). Here the identification of user is 3. Same SPVID should be assigned for the same network user on different PEs. When packet reaches PE1 from CE1, it carries the VLAN tag 200-300 of the user internal network. Since the dot1q-tunnel function is enabled, the user port on PE1 will add on the packet another VLAN tag, of which the ID is the SPVID assigned to the user. Afterwards, the packet will only be transmitted in VLAN3 when

traveling in the ISP internet network while carrying two VLAN tags (the inner tag is added when entering PE1, and the outer is SPVID), whereas the VLAN information of the user network is open to the provider network. When the packet reaches PE2 and before being forwarded to CE2 from the client port on PE2, the outer VLAN tag is removed, then the packet CE2 receives is absolutely identical to the one sent by CE1. For the user, the role the operator network plays between PE1 and PE2, is to provide a reliable layer-2 link.

The technology of Dot1q-tuunel provides the ISP internet the ability of supporting many client VLANs by only one VLAN of theirselves. Both the ISP internet and the clients can configure their own VLAN independently.

It is obvious that, the dot1q-tunnel function has got following characteristics:

Applicable through simple static configuration, no complex configuration or maintenance to be needed.

Operators will only have to assign one SPVID for each user, which increases the number of concurrent supportable users; while the users has got the ultimate freedom in selecting and managing the VLAN IDs (select within 1~4096 at users' will).

The user network is considerably independent. When the ISP internet is upgrading their network, the user networks do not have to change their original configuration.

Detailed description on the application and configuration of dot1q-tunnel will be provided in this section.

9.3.2 Dot1q-tunnel Configuration

Configuration Task Sequence of Dot1q-Tunnel:

1. Configure the dot1q-tunnel function on port
2. Configure the protocol type (TPID) on port

1. Configure the dot1q-tunnel function on port

Command	Explanation
Port mode	
dot1q-tunnel enable **no dot1q-tunnel enable**	Enter/exit the dot1q-tunnel mode on the port.

2. Configure the protocol type (TPID) on port

Command	Explanation
Port mode	
dot1q-tunnel tpid {0x8100|0x9100|0x9200|<1-65535>}	Configure the protocol type on TRUNK port.

9.3.3 Typical Applications of the Dot1q-tunnel

Scenario:

Edge switch PE1 and PE2 of the ISP internet forward the VLAN200~300 data between CE1 and CE2 of the client network with VLAN3. The port1 of PE1 is connected to CE1, port10 is connected to public network, the TPID of the connected equipment is 9100; port1 of PE2 is connected to CE2, port10 is connected to public network.

Configuration Item	Configuration Explanation
VLAN3	Port1 of PE1 and PE2.
dot1q-tunnel	Port1 of PE1 and PE2.
tpid	9100

Configuration procedure is as follows:

PE1:

```
Switch(config)#vlan 3
Switch(Config-Vlan3)#switchport interface ethernet 1/0/1
Switch(Config-Vlan3)#exit
Switch(Config)#interface ethernet 1/0/1
Switch(Config-Ethernet1/0/1)# dot1q-tunnel enable
Switch(Config-Ethernet1/0/1)# exit
Switch(Config)#interface ethernet 1/0/1
Switch(Config-Ethernet1/0/1)#switchport mode trunk
Switch(Config-Ethernet1/0/1)#dot1q-tunnel tpid 0x9100
Switch(Config-Ethernet1/0/1)#exit
Switch(Config)#
```

PE2:

```
Switch(config)#vlan 3
Switch(Config-Vlan3)#switchport interface ethernet 1/0/1
Switch(Config-Vlan3)#exit
Switch(Config)#interface ethernet 1/0/1
Switch(Config-Ethernet1/0/1)# dot1q-tunnel enable
Switch(Config-Ethernet1/0/1)# exit
Switch(Config)#interface ethernet 1/0/1
Switch(Config-Ethernet1/0/1)#switchport mode trunk
Switch(Config-Ethernet1/0/1)#dot1q-tunnel tpid 0x9100
Switch(Config-Ethernet1/0/1)#exit
Switch(Config)#
```

9.3.4 Dot1q-tunnel Troubleshooting

Enabling dot1q-tunnel on Trunk port will make the tag of the data packet unpredictable which is not required in the application. So it is not recommended to enable dot1q-tunnel on Trunk port.
Enabled with STP/MSTP is not supported.
Enabled with PVLAN is not supported.

9.4 VLAN-translation Configuration

9.4.1 Introduction to VLAN-translation

VLAN translation, as one can tell from the name, which translates the original VLAN ID to new VLAN ID according to the user requirements so to exchange data across different VLANs. VLAN translation is classified to ingress translation and egress translation, this switch only supports switchover of ingress for VLAN ID.
Application and configuration of VLAN translation will be explained in detail in this section.

9.4.2 VLAN-translation Configuration

Configuration task sequence of VLAN-translation:
1. Configure the VLAN-translation function on the port
2. Configure the VLAN-translation relations on the port
3. Configure whether the packet is dropped when checking VLAN-translation is failing
4. Show the related configuration of vlan-translation

1. Configure the VLAN-translation of the port

Command	Explanation
Port mode	
vlan-translation enable **no vlan-translation enable**	Enter/exit the port VLAN-translation mode.

2. Configure the VLAN-translation relation of the port

Command	Explanation
Port mode	
vlan-translation <*old-vlan-id*> to <*new-vlan-id*> in **no vlan-translation old-vlan-id in**	Add/delete a VLAN-translation relation.

3. Configure whether the packet is dropped when checking VLAN-translation is failing

Command	Explanation

Port mode	
vlan-translation miss drop in **no vlan-translation miss drop in**	Configure the VLAN-translation packet dropped on port if there is any failure.

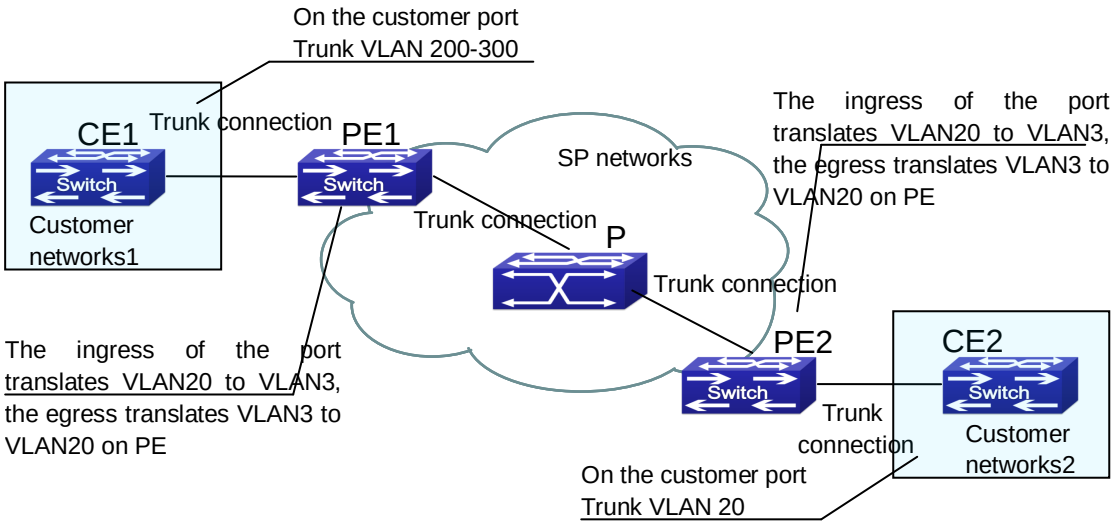
4. Show the related configuration of vlan-translation

Command	Explanation
Admin mode	
show vlan-translation	Show the related configuration of vlan-translation.

9.4.3 Typical application of VLAN-translation

Scenario:

Edge switch PE1 and PE2 of the ISP internet support the VLAN20 data task between CE1 and CE2 of the client network with VLAN3. The port1 of PE1 is connected to CE1, port10 is connected to public network; port1 of PE2 is connected to CE2, port10 is connected to public network.



Vlan translation topology mode

Configuration Item	Configuration Explanation
VLAN-translation	Port1 of PE1 and PE2.
Trunk port	Port1 and Port10 of PE1 and PE2.

Configuration procedure is as follows:

PE1, PE2:

switch(Config)#interface ethernet 1/0/1

switch(Config-Ethernet1/0/1)#switchport mode trunk

```
switch(Config-Ethernet1/0/1)# vlan-translation enable
switch(Config-Ethernet1/0/1)# vlan-translation 20 to 3 in
switch(Config-Ethernet1/0/1)# vlan-translation 3 to 20 out
switch(Config-Ethernet1/0/1)# exit
switch(Config)#interface ethernet 1/0/1
switch(Config-Ethernet1/0/1)#switchport mode trunk
switch(Config-Ethernet1/0/1)#exit
switch(Config)#
```

9.4.4 VLAN-translation Troubleshooting

Normally the VLAN-translation is applied on trunk ports.
Priority of vlan translation and vlan ingress filtering for processing packets is: vlan translation > vlan ingress filtering

9.5 Dynamic VLAN Configuration

9.5.1 Introduction to Dynamic VLAN

The dynamic VLAN is named corresponding to the static VLAN (namely the port based VLAN). Dynamic VLAN supported by the switch includes MAC-based VLAN, IP-subnet-based VLAN and Protocol-based VLAN. Detailed description is as follows:
The MAC-based VLAN division is based on the MAC address of each host, namely every host with a MAC address will be assigned to certain VLAN. By the means, the network user will maintain his membership in his belonging VLAN when moves from a physical location to another. As we can see the greatest advantage of this VLAN division is that the VLAN does not have to be re-configured when the user physic location change, namely shift from one switch to another, which is because it is user based, not switch port based.
The IP subnet based VLAN is divided according to the source IP address and its subnet mask of every host. It assigns corresponding VLAN ID to the data packet according to the subnet segment, leading the data packet to specified VLAN. Its advantage is the same as that of the MAC-based VLAN: the user does not have to change configuration when relocated.
The VLAN is divided by the network layer protocol, assigning different protocol to different VLANs. This is very attractive to the network administrators who wish to organize the user by applications and services. Moreover the user can move freely within the network while maintaining his membership. Advantage of this method enables user to change physical position without changing their VLAN residing configuration, while the VLAN can be divided by types of protocols which is important to the network administrators. Further, this method has no need of added frame label to identify the VLAN which reduce the network traffic.

Notice: Dynamic VLAN needs to associate with Hybrid attribute of the ports to work, so the ports that may be added to a dynamic VLAN must be configured as Hybrid port.

9.5.2 Dynamic VLAN Configuration

Dynamic VLAN Configuration Task Sequence:

1. Configure the MAC-based VLAN function on the port
2. Set the VLAN to MAC VLAN
3. Configure the correspondence between the MAC address and the VLAN
4. Configure the IP-subnet-based VLAN function on the port
5. Configure the correspondence between the IP subnet and the VLAN
6. Configure the correspondence between the Protocols and the VLAN
7. Adjust the priority of the dynamic VLAN

1. Configure the MAC-based VLAN function on the port

Command	Explanation
Port Mode	
switchport mac-vlan enable **no switchport mac-vlan enable**	Enable/disable the MAC-based VLAN function on the port.

2. Set the VLAN to MAC VLAN

Command	Explanation
Global Mode	
mac-vlan vlan <*vlan-id*> **no mac-vlan**	Configure the specified VLAN to MAC VLAN; the "**no mac-vlan**" command cancels the MAC VLAN configuration of this VLAN.

3. Configure the correspondence between the MAC address and the VLAN

Command	Explanation
Global Mode	
mac-vlan mac <mac-addrss> vlan <vlan-id> priority <priority-id> **no mac-vlan {mac <*mac-addrss*>|all}**	Add/delete the correspondence between the MAC address and the VLAN, namely specified MAC address join/leave specified VLAN.

4. Configure the IP-subnet-based VLAN function on the port

Command	Explanation

Port Mode	
switchport subnet-vlan enable **no switchport subnet-vlan enable**	Enable/disable the port IP-subnet-base VLAN function on the port.

5. Configure the correspondence between the IP subnet and the VLAN

Command	Explanation
Global Mode	
subnet-vlan ip-address <ipv4-addrss> mask <subnet-mask> vlan <vlan-id> priority <priority-id> **no subnet-vlan {ip-address *<ipv4-addrss>* mask <*subnet-mask*>|all}**	Add/delete the correspondence between the IP subnet and the VLAN, namely specified IP subnet joins/leaves specified VLAN.

6. Configure the correspondence between the Protocols and the VLAN

Command	Explanation
Global Mode	
protocol-vlan mode {ethernetii etype <*etype-id*>|llc {dsap <*dsap-id*> ssap <*ssap-id*>}|snap etype <*etype-id*>} vlan <*vlan-id*> priority <*priority-id*> **no protocol-vlan {mode {ethernetii etype <*etype-id*>|llc {dsap <*dsap-id*> ssap <*ssap-id*>}|snap etype <*etype-id*>}|all}**	Add/delete the correspondence between the Protocols and the VLAN, namely specified protocol joins/leaves specified VLAN.

7. Adjust the priority of the dynamic VLAN

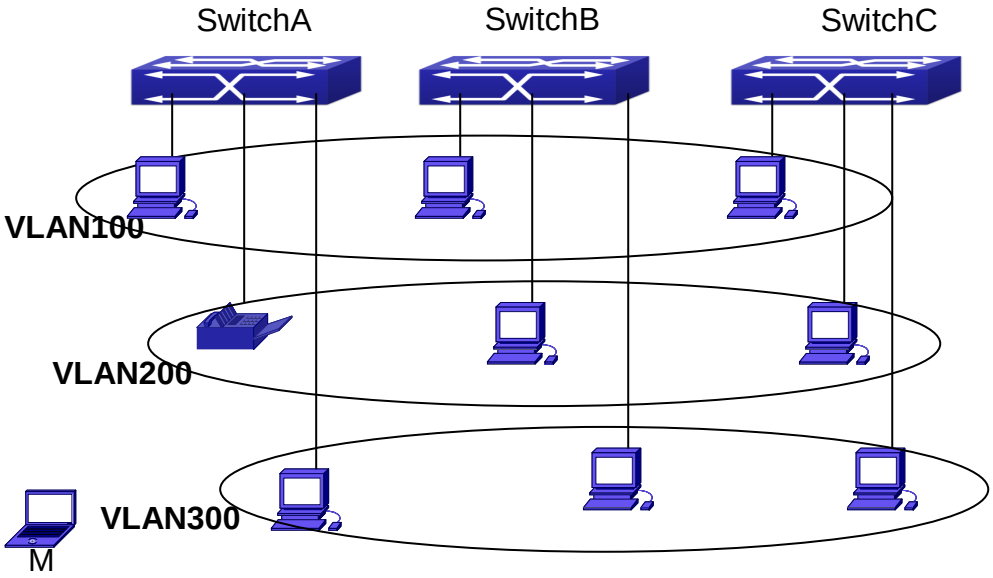
Command	Explanation
Global Mode	
dynamic-vlan mac-vlan prefer **dynamic-vlan subnet-vlan prefer**	Configure the priority of the dynamic VLAN.

9.5.3 Typical Application of the Dynamic VLAN

Scenario:

In the office network Department A belongs to VLAN100. Several members of this department often have the need to move within the whole office network. It is also required to ensure the resource for other members of the department to access VLAN 100. Assume one of the members is M, the MAC address of his PC is 00-1f-ce-11-22-33, when M moves to VLAN200

or VLAN300, the port connecting M is configured as Hybrid mode and belongs to VLAN100 with untag mode. In this way, the data of VLAN100 will be forwarded to the port connecting M, and implement the communication requirement in VLAN100.



Typical topology application of dynamic VLAN

Configuration Items	Configuration Explanation
MAC-based VLAN	Global configuration on Switch A, Switch B, Switch C.

For example, M at E1/0/1 of SwitchA, then the configuration procedures are as follows:
Switch A, Switch B, Switch C:
SwitchA (Config)#mac-vlan mac 00-1f-ce-11-22-33 vlan 100 priority 0
SwitchA (Config)#interface ethernet 1/0/1
SwitchA (Config-Ethernet1/0/1)# swportport mode hybrid
SwitchA (Config-Ethernet1/0/1)# swportport hybrid allowed vlan 100 untagged

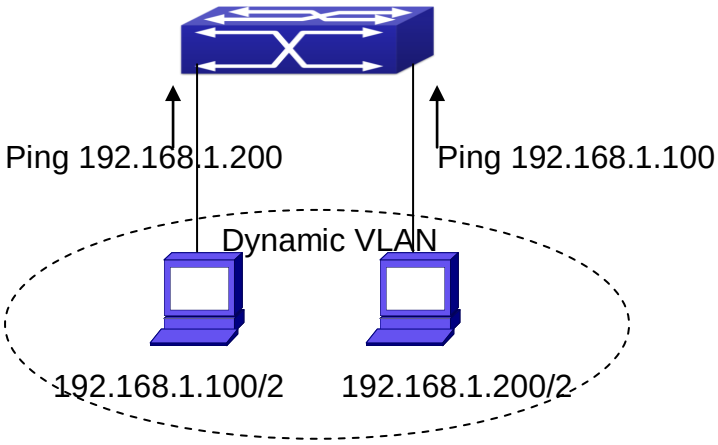
SwitchB (Config)#mac-vlan mac 00-1f-ce-11-22-33 vlan 100 priority 0
SwitchB (Config)#exit
SwitchB#

SwitchC (Config)#mac-vlan mac 00-1f-ce-11-22-33 vlan 100 priority 0
SwitchC (Config)#exit
SwitchC#

9.5.4 Dynamic VLAN Troubleshooting

On the switch configured with dynamic VLAN, if the two connected equipment (e.g. PC) are both belongs to the same dynamic VLAN, first communication between the two equipments

may not go through. The solution will be letting the two equipments positively send data packet to the switch (such as ping), to let the switch learn their source MAC, then the two equipments will be able to communicate freely within the dynamic VLAN.



Dynamic VLAN Troubleshooting

Priority of dynamic vlan and vlan ingress filtering for processing packets is: dynamic vlan> vlan ingress filtering

9.6 Voice VLAN Configuration

9.6.1 Introduction to Voice VLAN

Voice VLAN is specially configured for the user voice data traffic. By setting a Voice VLAN and adding the ports of the connected voice equipments to the Voice VLAN, the user will be able to configure QoS (Quality of service) service for voice data, and improve the voice data traffic transmission priority to ensure the calling quality.

The switch can judge if the data traffic is the voice data traffic from specified equipment according to the source MAC address field of the data packet entering the port. The packet with the source MAC address complying with the system defined voice equipment OUI (Organizationally Unique Identifier) will be considered the voice data traffic and transmitted to the Voice VLAN.

The configuration is based on MAC address, acquiring a mechanism in which every voice equipment transmitting information through the network has got its unique MAC address. VLAN will trace the address belongs to specified MAC. By This means, VLAN allows the voice equipment always belong to Voice VLAN when relocated physically. The greatest advantage of the VLAN is the equipment can be automatically placed into Voice VLAN according to its voice traffic which will be transmitted at specified priority. Meanwhile, when voice equipment is physically relocated, it still belongs to the Voice VLAN without any further configuration modification, which is because it is based on voice equipment other than switch port.

Notice: Voice VLAN needs to associate with Hybrid attribute of the ports to work, so the ports

that may be added to Voice VLAN must be configured as Hybrid port.

9.6.2 Voice VLAN Configuration

Voice VLAN Configuration Task Sequence:
Set the VLAN to Voice VLAN
Add a voice equipment to Voice VLAN
Enable the Voice VLAN on the port

1. Configure the VLAN to Voice VLAN

Command	Explanation
Global Mode	
voice-vlan vlan <*vlan-id*> **no voice-vlan**	Set/cancel the VLAN as a Voice VLAN

2. Add a Voice equipment to a Voice VLAN

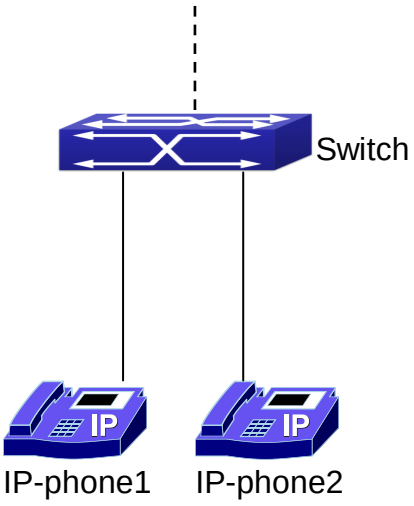
Command	Explanation
Global Mode	
voice-vlan mac <*mac-address*> mask <*mac-mask*> priority <*priority-id*> [name <*voice-name*>] **no voice-vlan {mac <*mac-address*> mask <*mac-mask*>|name <*voice-name*> |all}**	Specify certain voice equipment join/leave the Voice VLAN

3. Enable the Voice VLAN of the port

Command	Explanation
Port Mode	
switchport voice-vlan enable **no switchport voice-vlan enable**	Enable/disable the Voice VLAN function on the port

9.6.3 Typical Applications of the Voice VLAN

Scenario:
A company realizes voice communication through configuring Voice VLAN. IP-phone1 and IP-phone2 can be connected to any port of the switch, namely normal communication and interconnected with other switches through the uplink port. IP-phone1 MAC address is 00-1f-ce-11-22-33, connect port 1/0/1 of the switch, IP-phone2 MAC address is 00-1f-ce-11-22-55, connect port 1/0/2 of the switch.



VLAN typical apply topology Figure

Configuration items	Configuration Explanation
Voice VLAN	Global configuration on the Switch.

Configuration procedure:
Switch 1:
Switch(config)#vlan 100
Switch(Config-Vlan100)#exit
Switch(config)#voice-vlan vlan 100
Switch(config)#voice-vlan mac 00-1f-ce-11-22-33 mask 255 priority 5 name company
Switch(config)#voice-vlan mac 00-1f-ce-11-22-55 mask 255 priority 5 name company
Switch(config)#interface ethernet 1/0/10
Switch(Config-If-Ethernet1/0/10)#switchport mode trunk
Switch(Config-If-Ethernet1/0/10)#exit
switch(Config)#interface ethernet 1/0/1
switch(Config-If-Ethernet1/0/1)#switchport mode hybrid
switch(Config-If-Ethernet1/0/1)#switchport hybrid allowed vlan 100 untag
switch(Config-If-Ethernet1/0/1)#exit
switch(Config)#interface ethernet 1/0/2
switch(Config-If-Ethernet1/0/2)#switchport mode hybrid
switch(Config-If-Ethernet1/0/2)#switchport hybrid allowed vlan 100 untag
switch(Config-If-Ethernet1/0/2)#exit

9.6.4 Voice VLAN Troubleshooting

Voice VLAN can not be applied concurrently with MAC-base VLAN.
The Voice VLAN on the port is enabled by default. If the configured data can no longer enter the Voice VLAN during operation, please check if the Voice VLAN function has been disabled on the port.

Chapter 10 MAC Table Configuration

10.1 Introduction to MAC Table

MAC table is a table identifies the mapping relationship between destination MAC addresses and switch ports. MAC addresses can be categorized as static MAC addresses and dynamic MAC addresses. Static MAC addresses are manually configured by the user, have the highest priority and are permanently effective (will not be overwritten by dynamic MAC addresses); dynamic MAC addresses are entries learnt by the switch in data frame forwarding, and is effective for a limited period. When the switch receives a data frame to be forwarded, it stores the source MAC address of the data frame and creates a mapping to the destination port. Then the MAC table is queried for the destination MAC address, if hit, the data frame is forwarded in the associated port, otherwise, the switch forwards the data frame to its broadcast domain. If a dynamic MAC address is not learnt from the data frames to be forwarded for a long time, the entry will be deleted from the switch MAC table.

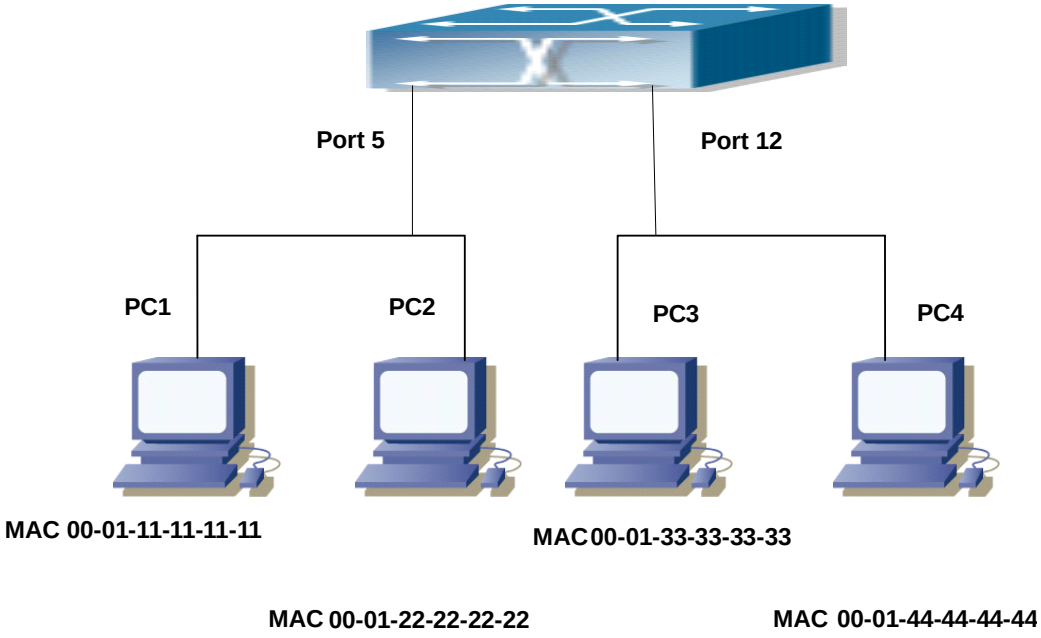
There are two MAC table operations:

Obtain a MAC address.

Forward or filter data frame according to the MAC table.

10.1.1 Obtaining MAC Table

The MAC table can be built up statically and dynamically. Static configuration is to set up a mapping between the MAC addresses and the ports; dynamic learning is the process in which the switch learns the mapping between MAC addresses and ports, and updates the MAC table regularly. In this section, we will focus on the dynamic learning process of MAC table.



MAC Table dynamic learning

The topology of the figure above: 4 PCs connected to switch, where PC1 and PC2 belongs to a same physical segment (same collision domain), the physical segment connects to port 1/0/5 of switch; PC3 and PC4 belongs to the same physical segment that connects to port 1/0/12 of switch.

The initial MAC table contains no address mapping entries. Take the communication of PC1 and PC3 as an example, the MAC address learning process is as follow:

When PC1 sends message to PC3, the switch receives the source MAC address 00-01-11-11-11-11 from this message, the mapping entry of 00-01-11-11-11-11 and port 1/0/5 is added to the switch MAC table.

At the same time, the switch learns the message is destined to 00-01-33-33-33-33, as the MAC table contains only a mapping entry of MAC address 00-01-11-11-11-11 and port1/0/5, and no port mapping for 00-01-33-33-33-33 present, the switch broadcast this message to all the ports in the switch (assuming all ports belong to the default VLAN1).

PC3 and PC4 on port 1/0/12 receive the message sent by PC1, but PC4 will not reply, as the destination MAC address is 00-01-33-33-33-33, only PC3 will reply to PC1. When port 1/0/12 receives the message sent by PC3, a mapping entry for MAC address 00-01-33-33-33-33 and port 1/0/12 is added to the MAC table.

Now the MAC table has two dynamic entries, MAC address 00-01-11-11-11-11 - port 1/0/5 and 00-01-33-33-33-33 -port1/0/12.

After the communication between PC1 and PC3, the switch does not receive any message sent from PC1 and PC3. And the MAC address mapping entries in the MAC table are deleted in 300 to 2*300 seconds (ie, in single to double aging time). The 300 seconds here is the default aging time for MAC address entry in switch. Aging time can be modified in switch.

10.1.2 Forward or Filter

The switch will forward or filter received data frames according to the MAC table. Take the above figure as an example, assuming switch have learnt the MAC address of PC1 and PC3, and the user manually configured the mapping relationship for PC2 and PC4 to ports. The MAC table of switch will be:

MAC Address	Port number	Entry added by
00-01-11-11-11-11	1/0/5	Dynamic learning
00-01-22-22-22-22	1/0/5	Static configuration
00-01-33-33-33-33	1/0/12	Dynamic learning
00-01-44-44-44-44	1/0/12	Static configuration

Forward data according to the MAC table

If PC1 sends a message to PC3, the switch will forward the data received on port 1/0/5 from port1/0/12.

Filter data according to the MAC table

If PC1 sends a message to PC2, the switch, on checking the MAC table, will find PC2 and

PC1 are in the same physical segment and filter the message (i.e. drop this message).
Three types of frames can be forwarded by the switch:
Broadcast frame
Multicast frame
Unicast frame
The following describes how the switch deals with all the three types of frames:
Broadcast frame: The switch can segregate collision domains but not broadcast domains. If no VLAN is set, all devices connected to the switch are in the same broadcast domain. When the switch receives a broadcast frame, it forwards the frame in all ports. When VLANs are configured in the switch, the MAC table will be adapted accordingly to add VLAN information. In this case, the switch will not forward the received broadcast frames in all ports, but forward the frames in all ports in the same VLAN.
Multicast frame: For the unknown multicast, the switch will broadcast it in the same vlan, but the switch only forwards the multicast frames to the multicast group's port if IGMP Snooping function or the static multicast group has been configured.
Unicast frame: When no VLAN is configured, if the destination MAC addresses are in the switch MAC table, the switch will directly forward the frames to the associated ports; when the destination MAC address in a unicast frame is not found in the MAC table, the switch will broadcast the unicast frame. When VLANs are configured, the switch will forward unicast frame within the same VLAN. If the destination MAC address is found in the MAC table but belonging to different VLANs, the switch can only broadcast the unicast frame in the VLAN it belongs to.

10.2 Mac Address Table Configuration Task List

Configure the MAC address aging-time
Configure static MAC forwarding or filter entry
Clear dynamic address table

Configure the MAC aging-time

Command	Explanation
Global Mode	
mac-address-table aging-time <*0*|*aging-time*> **no mac-address-table aging-time**	Configure the MAC address aging-time.

Configure static MAC forwarding or filter entry

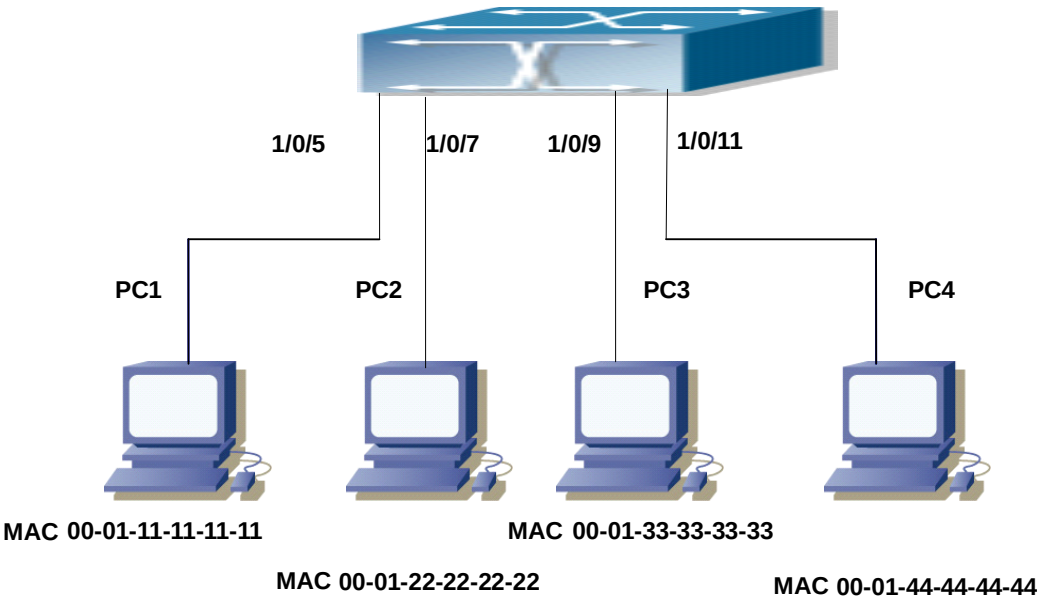
Command	Explanation
Global Mode	
**mac-address-table {static | static-multicast | blackhole} address <*mac-addr*> vlan <*vlan-*	Configure static MAC entries, static multicast MAC entries, filter address

***id* > [interface [ethernet | portchannel] *<interface-name>*] | [source|destination|both]** **no mac-address-table {static | static-multicast | blackhole | dynamic} [address *<mac-addr>*] [vlan *<vlan-id>*] [interface [ethernet | portchannel] *<interface-name>*]**	entires.

Clear dynamic address table

Command	Explanation
Admin Mode	
clear mac-address-table dynamic [address *<mac-addr>*] [vlan *<vlan-id>*] [interface [ethernet | portchannel] *<interface-name>*]	Clear the dynamic address table.

10.3 Typical Configuration Examples



MAC Table typical configuration example

Scenario:

Four PCs as shown in the above figure connect to port 1/0/5, 1/0/7, 1/0/9, 1/0/11 of switch, all the four PCs belong to the default VLAN1. As required by the network environment, dynamic learning is enabled. PC1 holds sensitive data and can not be accessed by any other PC that is in another physical segment; PC2 and PC3 have static mapping set to port 7 and port 9, respectively.

The configuration steps are listed below:

1. Set the MAC address 00-01-11-11-11-11 of PC1 as a filter address.

Switch(config)#mac-address-table static 00-01-11-11-11-11 discard vlan 1.
2.Set the static mapping relationship for PC2 and PC3 to port 7 and port 9, respectively.
Switch(config)#mac-address-table static address 00-01-22-22-22-22 vlan 1 interface ethernet 1/0/7
Switch(config)#mac-address-table static address 00-01-33-33-33-33 vlan 1 interface ethernet 1/0/9

10.4 MAC Table Troubleshooting

Using the show mac-address-table command, a port is found to be failed to learn the MAC of a device connected to it. Possible reasons:
The connected cable is broken.
Spanning Tree is enabled and the port is in “discarding” status; or the device is just connected to the port and Spanning Tree is still under calculation, wait until the Spanning Tree calculation finishes, and the port will learn the MAC address.
If not the problems mentioned above, please check for the switch portand contact technical support for solution.

10.5 MAC Address Function Extension

10.5.1 MAC Address Binding

10.5.1.1 Introduction to MAC Address Binding

Most switches support MAC address learning, each port can dynamically learn several MAC addresses, so that forwarding data streams between known MAC addresses within the ports can be achieved. If a MAC address is aged, the packet destined for that entry will be broadcasted. In other words, a MAC address learned in a port will be used for forwarding in that port, if the connection is changed to another port, the switch will learn the MAC address again to forward data in the new port.
However, in some cases, security or management policy may require MAC addresses to be bound with the ports, only data stream from the binding MAC are allowed to be forwarded in the ports. That is to say, after a MAC address is bound to a port, only the data stream destined for that MAC address can flow in from the binding port, data stream destined for the other MAC addresses that not bound to the port will not be allowed to pass through the port.

10.5.1.2 MAC Address Binding Configuration Task List

1. Enable MAC address binding function for the ports
2. Lock the MAC addresses for a port
3. MAC address binding property configuration
4. mac-notification trap configuration

1. Enable MAC address binding function for the ports

Command	Explanation
Port Mode	
switchport port-security **no switchport port-security**	Enable MAC address binding function for the port and lock the port. When a port is locked, the MAC address learning function for the port will be disabled: the "**no switchport port-security**" command disables the MAC address binding function for the port**,** and restores the MAC address learning function for the port.

2. Lock the MAC addresses for a port

Command	Explanation
Port Mode	
switchport port-security lock **no switchport port-security lock**	Lock the port, then MAC addresses learned will be disabled. The "**no switchport port-security lock**" command restores the function.
switchport port-security convert	Convert dynamic secure MAC addresses learned by the port to static secure MAC addresses.
switchport port-security timeout <*value*> **no switchport port-security timeout**	Enable port locking timer function; the "**no switchport port-security timeout**" restores the default setting.
switchport port-security mac-address <*mac-address*> **no switchport port-security mac-address <*mac-address*>**	Add static secure MAC address; the "**no switchport port-security mac-address**" command deletes static secure MAC address.
Admin Mode	
clear port-security dynamic [address <*mac-addr*> | interface <*interface-id*>]	Clear dynamic MAC addresses learned by the specified port.

3. MAC address binding property configuration

Command	Explanation
Port Mode	
switchport port-security maximum <*value*> **no switchport port-security maximum**	Set the maximum number of secure MAC addresses for a port; the "**no switchport**

<value>	**port-security maximum**” command restores the default value.
switchport port-security violation {protect | shutdown} [recovery <30-3600>] **no switchport port-security violation**	Set the violation mode for the port; the “**no switchport port-security violation**” command restores the default setting.

4. mac-notification trap configuration

Command	Explanation
Global Mode	
mac-address-table periodic-monitor-time <5-86400>	Set the MAC monitor interval to count the added and deleted MAC in time, and send out them with trap message.

10.5.1.3 Binding MAC Address Binding Troubleshooting

Enabling MAC address binding for ports may fail in some occasions. Here are some possible causes and solutions:

If MAC address binding cannot be enabled for a port, make sure the port is not enabling port aggregation and is not configured as a Trunk port. MAC address binding is exclusive to such configurations. If MAC address binding is to be enabled, the functions mentioned above must be disabled first.

If a secure address is set as static address and deleted, that secure address will be unusable even though it exists. For this reason, it is recommended to avoid static address for ports enabling MAC address.

Chapter 11 MSTP Configuration

11.1 Introduction to MSTP

The MSTP (Multiple STP) is a new spanning-tree protocol which is based on the STP and the RSTP. It runs on all the bridges of a bridged-LAN. It calculates a common and internal spanning tree (CIST) for the bridge-LAN which consists of the bridges running the MSTP, the RSTP and the STP. It also calculates the independent multiple spanning-tree instances (MSTI) for each MST domain (MSTP domain). The MSTP, which adopts the RSTP for its rapid convergence of the spanning tree, enables multiple VLANs to be mapped to the same spanning-tree instance which is independent to other spanning-tree instances. The MSTP provides multiple forwarding paths for data traffic and enables load balancing. Moreover, because multiple VLANs share a same MSTI, the MSTP can reduce the number of spanning-tree instances, which consumes less CPU resources and reduces the bandwidth consumption.

11.1.1 MSTP Region

Because multiple VLANs can be mapped to a single spanning tree instance, IEEE 802.1s committee raises the MST concept. The MST is used to make the association of a certain VLAN to a certain spanning tree instance.

A MSTP region is composed of one or multiple bridges with the same MCID (MST Configuration Identification) and the bridged-LAN (a certain bridge in the MSTP region is the designated bridge of the LAN, and the bridges attaching to the LAN are not running STP). All the bridges in the same MSTP region have the same MSID.

MSID consists of 3 attributes:

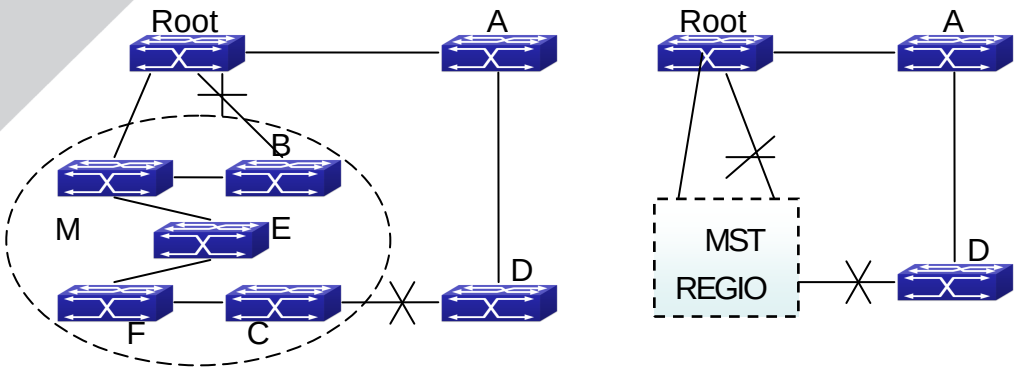
Configuration Name: Composed by digits and letters

Revision Level

Configuration Digest: VLANs mapping to spanning tree instances

The bridges with the same 3 above attributes are considered as in the same MST region.

When the MSTP calculates CIST in a bridged-LAN, a MSTP region is considered as a bridge. See the figure below:



Example of CIST and MST Region

In the above network, if the bridges are running the STP or the RSTP, one port between Bridge M and Bridge B should be blocked. But if the bridges in the yellow range run the MSTP and are configured in the same MST region, MSTP will treat this region as a bridge. Therefore, one port between Bridge B and Root is blocked and one port on Bridge D is blocked.

11.1.1.1 Operations within an MSTP Region

The IST connects all the MSTP bridges in a region. When the IST converges, the root of the IST becomes the IST master, which is the switch within the region with the lowest bridge ID and path cost to the CST root. The IST master is also the CST root if there is only one region within the network. If the CST root is outside the region, one of the MSTP bridges at the boundary of the region is selected as the IST master.

When an MSTP bridge initializes, it sends BPDUs claiming itself as the root of the CST and the IST master, with both of the path costs to the CST root and to the IST master set to zero. The bridge also initializes all of its MST instances and claims to be the root for all of them. If the bridge receives superior MST root information (lower bridge ID, lower path cost, and so forth) than currently stored for the port, it relinquishes its claim as the IST master.

Within a MST region, the IST is the only spanning-tree instance that sends and receives BPDUs. Because the MST BPDU carries information for all instances, the number of BPDUs that need to be processed by a switch to support multiple spanning-tree instances is significantly reduced.

All MST instances within the same region share the same protocol timers, but each MST instance has its own topology parameters, such as root switch ID, root path cost, and so forth.

11.1.1.2 Operations between MST Regions

If there are multiple regions or legacy 802.1D bridges within the network, MSTP establishes and maintains the CST, which includes all MST regions and all legacy STP bridges in the network. The MST instances combine with the IST at the boundary of the region to become the CST.

The MSTI is only valid within its MST region. An MSTI has nothing to do with MSTIs in other MST regions. The bridges in a MST region receive the MST BPDU of other regions through

Boundary Ports. They only process CIST related information and abandon MSTI information.

11.1.2 Port Roles

The MSTP bridge assigns a port role to each port which runs MSTP.
CIST port roles: Root Port, Designated Port, Alternate Port and Backup Port
On top of those roles, each MSTI port has one new role: Master Port.
The port roles in the CIST (Root Port, Designated Port, Alternate Port and Backup Port) are defined in the same ways as those in the RSTP.

11.1.3 MSTP Load Balance

In a MSTP region, VLANs can by mapped to various instances. That can form various topologies. Each instance is independent from the others and each distance can have its own attributes such as bridge priority and port cost etc. Consequently, the VLANs in different instances have their own paths. The traffic of the VLANs are load-balanced.

11.2 MSTP Configuration Task List

MSTP configuration task list:
1. Enable the MSTP and set the running mode
2. Configure instance parameters
3. Configure MSTP region parameters
4. Configure MSTP time parameters
5. Configure the fast migrate feature for MSTP
6. Configure the format of port packet
7. Configure the spanning-tree attribute of port
8. Configure the snooping attribute of authentication key
9. Configure the FLUSH mode once topology changes

1. Enable MSTP and set the running mode

Command	Explanation
Global Mode and Port Mode	
spanning-tree **no spanning-tree**	Enable/Disable MSTP.
Global Mode	
spanning-tree mode {mstp|stp|rstp} **no spanning-tree mode**	Set MSTP running mode.
Port Mode	
spanning-tree mcheck	Force port migrate to run under MSTP.

2. Configure instance parameters

Command	Explanation
Global Mode	
spanning-tree mst <instance-id> priority <bridge-priority> **no spanning-tree mst *<instance-id>* priority**	Set bridge priority for specified instance.
spanning-tree priority *<bridge-priority>* **no spanning-tree priority**	Configure the spanning-tree priority of the switch.
Port Mode	
spanning-tree mst *<instance-id>* cost *<cost>* **no spanning-tree mst *<instance-id>* cost**	Set port path cost for specified instance.
spanning-tree mst *<instance-id>* port-priority *<port-priority>* **no spanning-tree mst *<instance-id>* port-priority**	Set port priority for specified instance.
spanning-tree mst *<instance-id>* rootguard **no spanning-tree mst *<instance-id>* rootguard**	Configure currently port whether running rootguard in specified instance, configure the rootguard port can't turn to root port.
spanning-tree rootguard **no spanning-tree rootguard**	Configure currently port whether running rootguard in instance 0, configure the rootguard port can't turn to root port.
spanning-tree [mst <instance-id>] loopguard **no spanning-tree [mst <instance-id>] loopguard**	Enable loopguard function on specified instance, the no command disables this function.

3. Configure MSTP region parameters

Command	Explanation
Global Mode	
spanning-tree mst configuration **no spanning-tree mst configuration**	Enter MSTP region mode. The no command restores the default setting.
MSTP region mode	
show	Display the information of the current running system.
instance <instance-id> vlan <vlan-list> **no instance <instance-id> [vlan <vlan-**	Create Instance and set mapping between VLAN and Instance.

list>]	
name <name> **no name**	Set MSTP region name.
revision-level <*level*> **no revision-level**	Set MSTP region revision level.
abort	Quit MSTP region mode and return to Global mode without saving MSTP region configuration.
exit	Quit MSTP region mode and return to Global mode with saving MSTP region configuration.
no	Cancel one command or set initial value.

4. Configure MSTP time parameters

Command	Explanation
Global Mode	
spanning-tree forward-time <*time*> **no spanning-tree forward-time**	Set the value for switch forward delay time.
spanning-tree hello-time <*time*> **no spanning-tree hello-time**	Set the Hello time for sending BPDU messages.
spanning-tree maxage <*time*> **no spanning-tree maxage**	Set Aging time for BPDU messages.
spanning-tree max-hop <*hop-count*> **no spanning-tree max-hop**	Set Maximum number of hops of BPDU messages in the MSTP region.

5. Configure the fast migrate feature for MSTP

Command	Explanation
Port Mode	
spanning-tree link-type p2p {auto|force-true|force-false} **no spanning-tree link-type**	Set the port link type.
spanning-tree portfast [bpdufilter| bpduguard] [recovery <30-3600>] **no spanning-tree portfast**	Set and cancel the port to be an boundary port. bpdufilter receives the BPDU discarding; bpduguard receives the BPDU will disable port; no parameter receives the BPDU, the port becomes a non-boundary port.

6. Configure the format of MSTP

Command	Explanation
Port Mode	

spanning-tree format standard **spanning-tree format privacy** **spanning-tree format auto** **no spanning-tree format**	Configure the format of port spanning-tree packet, standard format is provided by IEEE, privacy is compatible with CISCO and auto means the format is determined by checking the received packet.

7. Configure the spanning-tree attribute of port

Command	Explanation
Port Mode	
spanning-tree cost **no spanning-tree cost**	Set the port path cost.
spanning-tree port-priority **no spanning-tree port-priority**	Set the port priority.
spanning-tree rootguard **no spanning-tree rootguard**	Set the port is root port.
Global Mode	
spanning-tree transmit-hold-count <tx-hold-count-value> **no spanning-tree transmit-hold-count**	Set the max transmit-hold-count of port.

8. Configure the snooping attribute of authentication key

Command	Explanation
Port Mode	
spanning-tree digest-snooping **no spanning-tree digest-snooping**	Set the port to use the authentication string of partner port. The no command restores to use the generated string.

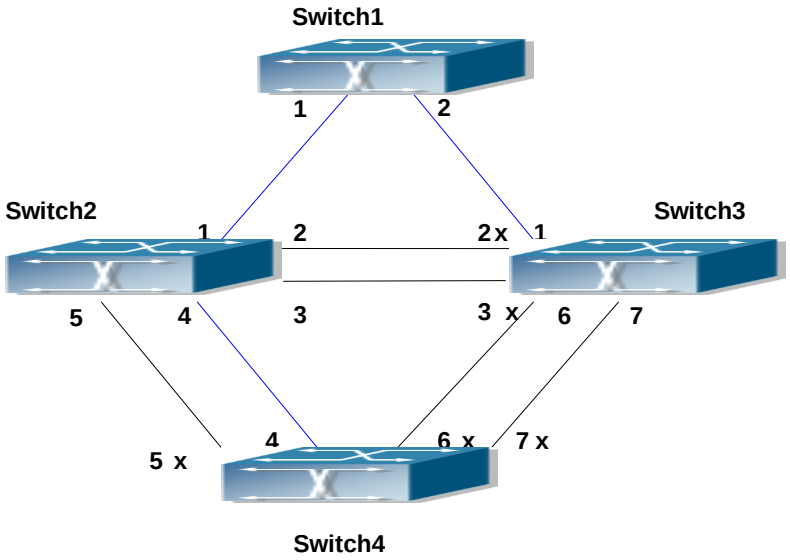
9. Configure the FLUSH mode once topology changes

Command	Explanation
Global Mode	
spanning-tree tcflush {enable| disable| protect} **no spanning-tree tcflush**	Enable: the spanning-tree flush once the topology changes. Disable: the spanning tree don't flush when the topology changes. Protect: the spanning-tree flush not more than one time every ten seconds. The no command restores to default setting, enable flush once the topology changes.
Port Mode	

spanning-tree tcflush {enable| disable| protect} **no spanning-tree tcflush**	Configure the port flush mode. The no command restores to use the global configured flush mode.

11.3 MSTP Example

The following is a typical MSTP application example:



Typical MSTP Application Scenario

The connections among the switches are shown in the above figure. All the switches run in the MSTP mode by default, their bridge priority, port priority and port route cost are all in the default values (equal). The default configuration for switches is listed below:

Bridge Name		Switch1	Switch2	Switch3	Switch4
Bridge MAC Address		…00-00-01	…00-00-02	…00-00-03	…00-00-04
Bridge Priority		32768	32768	32768	32768
Port Priority	Port 1	128	128	128	
	Port 2	128	128	128	
	Port 3		128	128	
	Port 4		128		128
	Port 5		128		128
	Port 6			128	128
	Port 7			128	128
Route	Port 1	200000	200000	200000	

Cost	Port 2	200000	200000	200000	
	Port 3		200000	200000	
	Port 4		200000		200000
	Port 5		200000		200000
	Port 6			200000	200000
	Port 7			200000	200000

By default, the MSTP establishes a tree topology (in blue lines) rooted with SwitchA. The ports marked with “x” are in the discarding status, and the other ports are in the forwarding status.

Configurations Steps:
Step 1: Configure port to VLAN mapping:
Create VLAN 20, 30, 40, 50 in Switch2, Switch3 and Switch4.
Set ports 1-7 as trunk ports in Switch2 Switch3 and Switch4.
Step 2: Set Switch2, Switch3 and Switch4 in the same MSTP:
Set Switch2, Switch3 and Switch4 to have the same region name as mstp.
Map VLAN 20 and VLAN 30 in Switch2, Switch3 and Switch4 to Instance 3; Map VLAN 40 and VLAN 50 in Switch2, Switch3 and Switch4 to Instance 4.
Step 3: Set Switch3 as the root bridge of Instance 3; Set Switch4 as the root bridge of Instance 4
Set the bridge priority of Instance 3 in Switch3 as 0.
Set the bridge priority of Instance 4 in Switch4 as 0.

The detailed configuration is listed below:

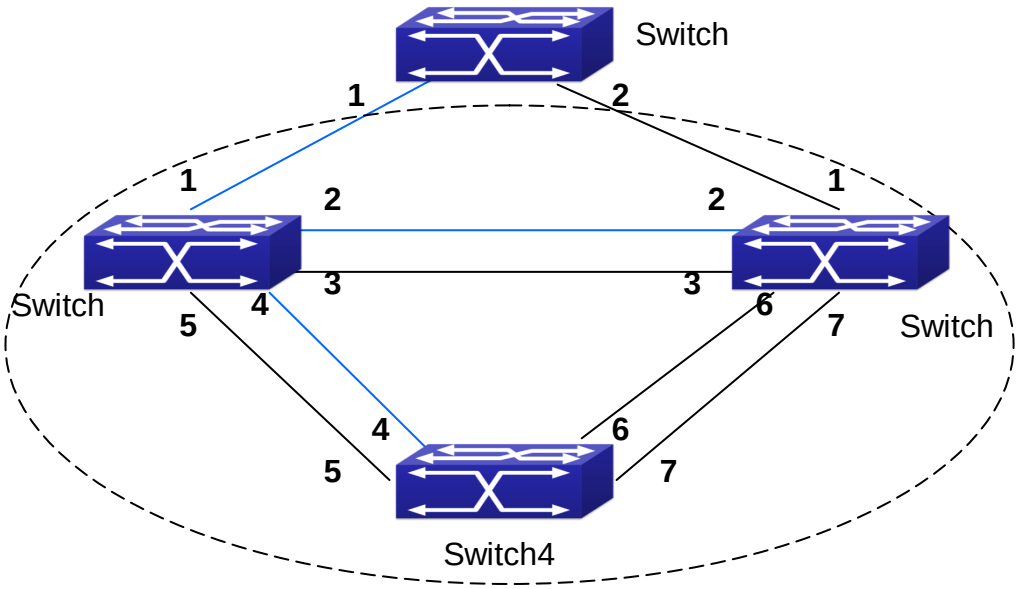
Switch2:
Switch2(config)#vlan 20
Switch2(Config-Vlan20)#exit
Switch2(config)#vlan 30
Switch2(Config-Vlan30)#exit
Switch2(config)#vlan 40
Switch2(Config-Vlan40)#exit
Switch2(config)#vlan 50
Switch2(Config-Vlan50)#exit
Switch2(config)#spanning-tree mst configuration
Switch2(Config-Mstp-Region)#name mstp
Switch2(Config-Mstp-Region)#instance 3 vlan 20;30
Switch2(Config-Mstp-Region)#instance 4 vlan 40;50
Switch2(Config-Mstp-Region)#exit
Switch2(config)#interface e1/0/1-7

```
Switch2(Config-Port-Range)#switchport mode trunk
Switch2(Config-Port-Range)#exit
Switch2(config)#spanning-tree

Switch3:
Switch3(config)#vlan 20
Switch3(Config-Vlan20)#exit
Switch3(config)#vlan 30
Switch3(Config-Vlan30)#exit
Switch3(config)#vlan 40
Switch3(Config-Vlan40)#exit
Switch3(config)#vlan 50
Switch3(Config-Vlan50)#exit
Switch3(config)#spanning-tree mst configuration
Switch3(Config-Mstp-Region)#name mstp
Switch3(Config-Mstp-Region)#instance 3 vlan 20;30
Switch3(Config-Mstp-Region)#instance 4 vlan 40;50
Switch3(Config-Mstp-Region)#exit
Switch3(config)#interface e1/0/1-7
Switch3(Config-Port-Range)#switchport mode trunk
Switch3(Config-Port-Range)#exit
Switch3(config)#spanning-tree
Switch3(config)#spanning-tree mst 3 priority 0

Switch4:
Switch4(config)#vlan 20
Switch4(Config-Vlan20)#exit
Switch4(config)#vlan 30
Switch4(Config-Vlan30)#exit
Switch4(config)#vlan 40
Switch4(Config-Vlan40)#exit
Switch4(config)#vlan 50
Switch4(Config-Vlan50)#exit
Switch4(config)#spanning-tree mst configuration
Switch4(Config-Mstp-Region)#name mstp
Switch4(Config-Mstp-Region)#instance 3 vlan 20;30
Switch4(Config-Mstp-Region)#instance 4 vlan 40;50
Switch4(Config-Mstp-Region)#exit
Switch4(config)#interface e1/0/1-7
Switch4(Config-Port-Range)#switchport mode trunk
```

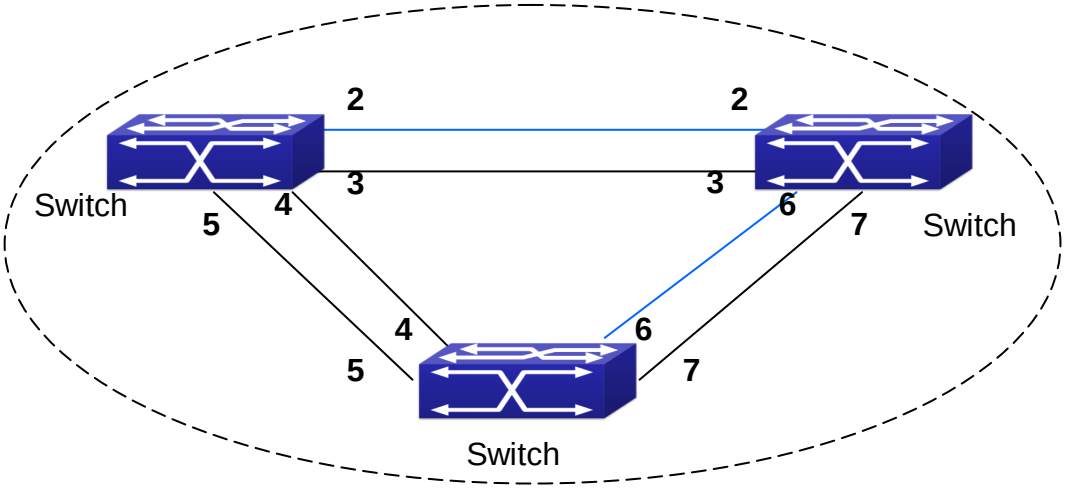
Switch4(Config-Port-Range)#exit
Switch4(config)#spanning-tree
Switch4(config)#spanning-tree mst 4 priority 0

After the above configuration, Switch1 is the root bridge of the instance 0 of the entire network. In the MSTP region which Switch2, Switch3 and Switch4 belong to, Switch2 is the region root of the instance 0, Switch3 is the region root of the instance 3 and Switch4 is the region root of the instance 4. The traffic of VLAN 20 and VLAN 30 is sent through the topology of the instance 3. The traffic of VLAN 40 and VLAN 50 is sent through the topology of the instance 4. And the traffic of other VLANs is sent through the topology of the instance 0. The port 1 in Switch2 is the master port of the instance 3 and the instance 4.

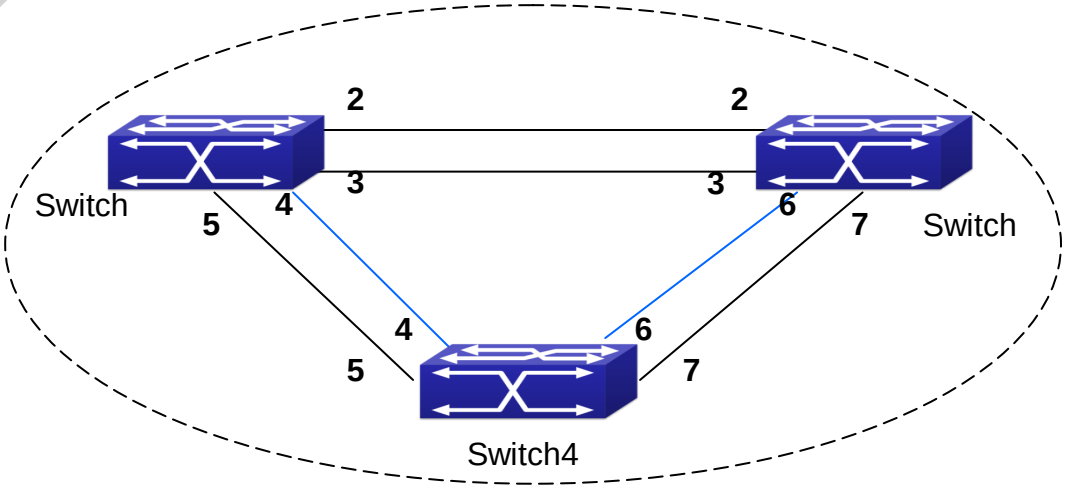
The MSTP calculation generates 3 topologies: the instance 0, the instance 3 and the instance 4 (marked with blue lines). The ports with the mark “x” are in the status of discarding. The other ports are the status of forwarding. Because the instance 3 and the instance 4 are only valid in the MSTP region, the following figure only shows the topology of the MSTP region.



The Topology Of the Instance 0 after the MSTP Calculation



The Topology Of the Instance 3 after the MSTP Calculation



The Topology Of the Instance 4 after the MSTP Calculation

11.4 MSTP Troubleshooting

In order to run the MSTP on the switch port, the MSTP has to be enabled globally. If the MSTP is not enabled globally, it can't be enabled on the port.

The MSTP parameters co work with each other, so the parameters should meet the following conditions. Otherwise, the MSTP may work incorrectly.

2×(Bridge_Forward_Delay -1.0 seconds) >= Bridge_Max_Age

Bridge_Max_Age >= 2 ×(Bridge_Hello_Time + 1.0 seconds)

When users modify the MSTP parameters, they have to be sure about the changes of the topologies. The global configuration is based on the bridge. Other configurations are based on the individual instances.

Chapter 12 QoS Configuration

12.1 Introduction to QoS

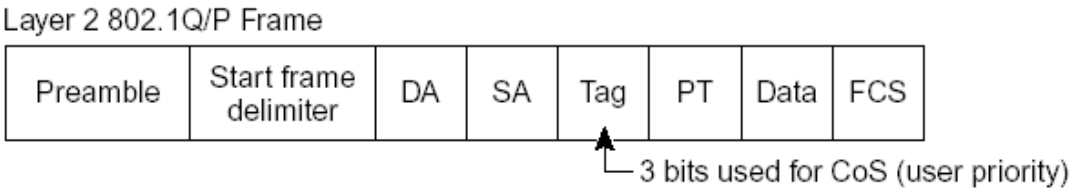
QoS (Quality of Service) is a set of capabilities that allow you to create differentiated services for network traffic, thereby providing better service for selected network traffic. QoS is a guarantee for service quality of consistent and predictable data transfer service to fulfill program requirements. QoS cannot generate extra bandwidth but provides more effective bandwidth management according to the application requirement and network management policy.

12.1.1 QoS Terms

QoS: Quality of Service, provides a guarantee for service quality of consistent and predictable data transfer service to fulfill program requirements. QoS cannot generate new bandwidth but provides more effective bandwidth management according to the application requirement and network management.

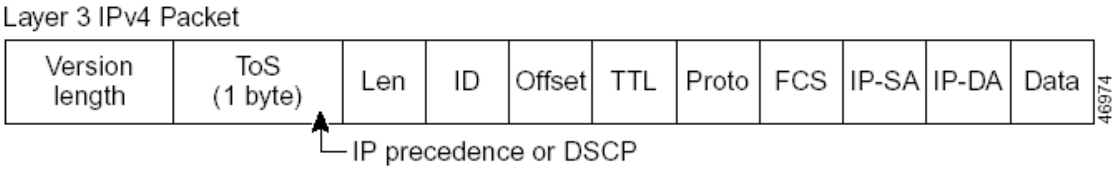
QoS Domain: QoS Domain supports QoS devices to form a net-topology that provides Quality of Service, so this topology is defined as QoS Domain.

CoS: Class of Service, the classification information carried by Layer 2 802.1Q frames, taking 3 bits of the Tag field in frame header, is called user priority level in the range of 0 to 7.



CoS priority

ToS: Type of Service, a one-byte field carried in Layer 3 IPv4 packet header to symbolize the service type of IP packets. Among ToS field can be IP Precedence value or DSCP value.



ToS priority

IP Precedence: IP priority. Classification information carried in Layer 3 IP packet header, occupying 3 bits, in the range of 0 to 7.

DSCP: Differentiated Services Code Point, classification information carried in Layer 3 IP packet header, occupying 6 bits, in the range of 0 to 63, and is downward compatible with IP Precedence.

Internal Priority: The internal priority setting of the switch chip, it's valid range relates with the chip, it's shortening is Int-Prio or IntP.
Drop Precedence: When processing the packets, firstly drop the packets with the bigger drop precedence, the ranging is 0-1. It's shortening is Drop-Prec or DP.
Classification: The entry action of QoS, classifying packet traffic according to the classification information carried in the packet and ACLs.
Policing: Ingress action of QoS that lays down the policing policy and manages the classified packets.
Remark: Ingress action of QoS, perform allowing, degrading or discarding operations to packets according to the policing policies.
Scheduling: QoS egress action. Add the packets to the corresponding egress queue according to the internal priority. And then decide sending and dropping according to Drop Precedence, sending algorithm and queue weight of egress queue.

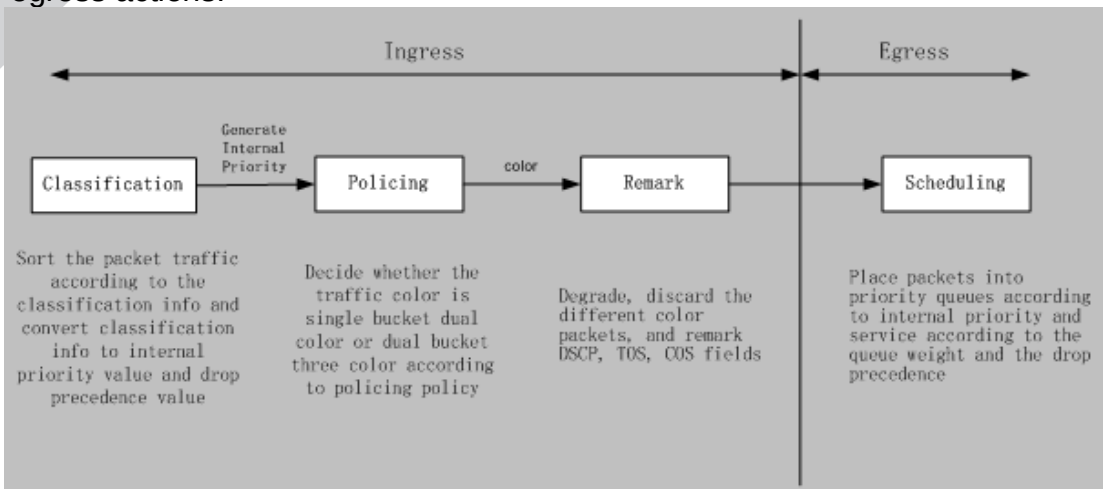
12.1.2 QoS Implementation

To implement the switch software QoS, a general, mature reference model should be given. QoS can not create new bandwidth, but can maximize the adjustment and configuration for the current bandwidth resource. Fully implemented QoS can achieve complete management over the network traffic. The following is as accurate as possible a description of QoS.
The data transfer specifications of IP cover only addresses and services of source and destination, and ensure correct packet transmission using OSI layer 4 or above protocols such as TCP. However, rather than provide a mechanism for providing and protecting packet transmission bandwidth, IP provide bandwidth service by the best effort. This is acceptable for services like Mail and FTP, but for increasing multimedia business data and e-business data transmission, this best effort method cannot satisfy the bandwidth and low-lag requirement.
Based on differentiated service, QoS specifies a priority for each packet at the ingress. The classification information is carried in Layer 3 IP packet header or Layer 2 802.1Q frame header. QoS provides same service to packets of the same priority, while offers different operations for packets of different priority. QoS-enabled switch or router can provide different bandwidth according to the packet classification information, and can remark on the classification information according to the policing policies configured, and may discard some low priority packets in case of bandwidth shortage.
If devices of each hop in a network support differentiated service, an end-to-end QoS solution can be created. QoS configuration is flexible, the complexity or simplicity depends on the network topology and devices and analysis to incoming/outgoing traffic.

12.1.3 Basic QoS Model

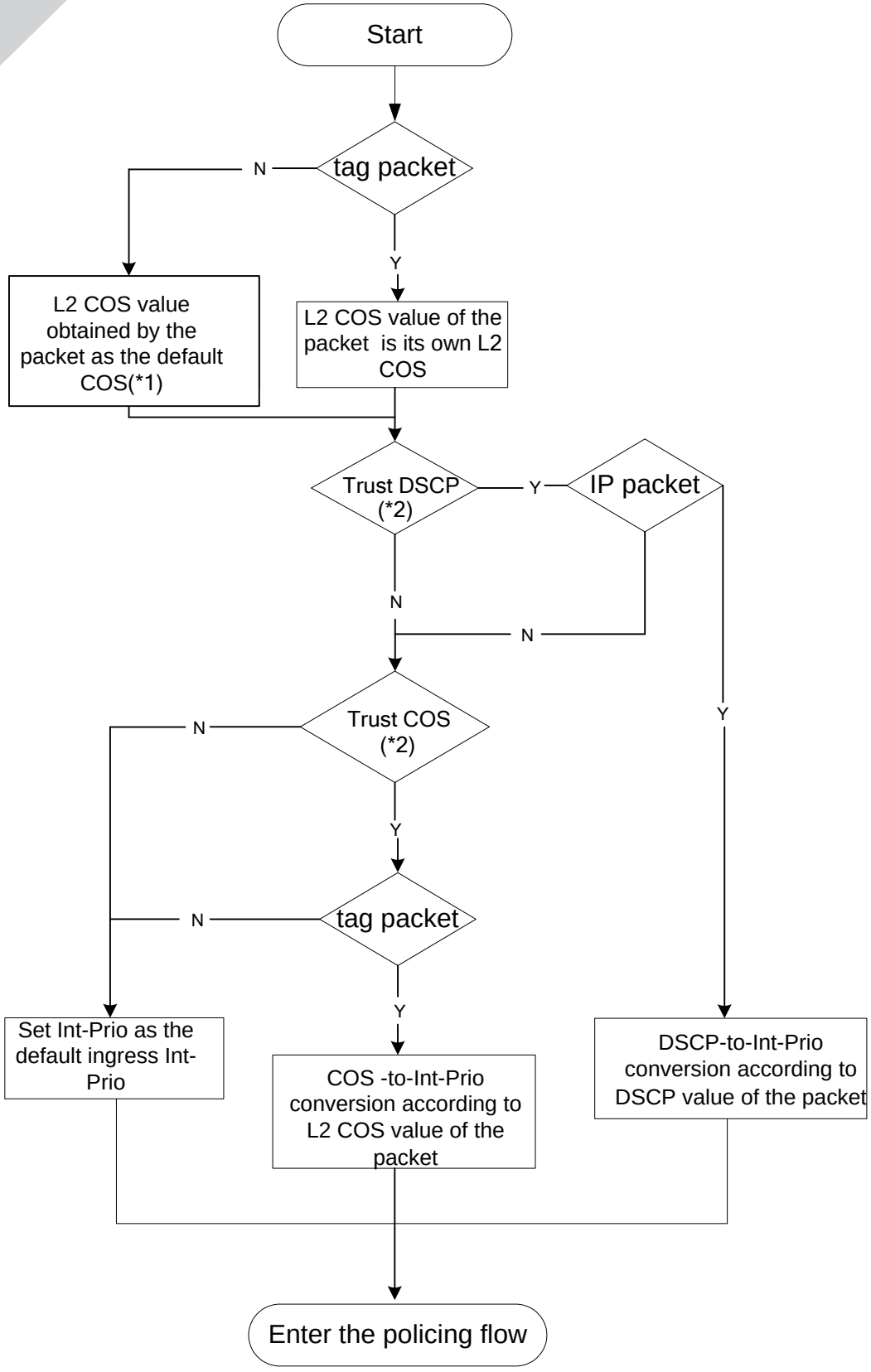
The basic QoS consists of four parts: Classification, Policing, Remark and Scheduling, where classification, policing and remark are sequential ingress actions, and Queuing and Scheduling

are QoS egress actions.



Basic QoS Model

Classification: Classify traffic according to packet classification information and generate internal priority based the classification information. For different packet types, classification is performed differently; the flowchart below explains this in detail.



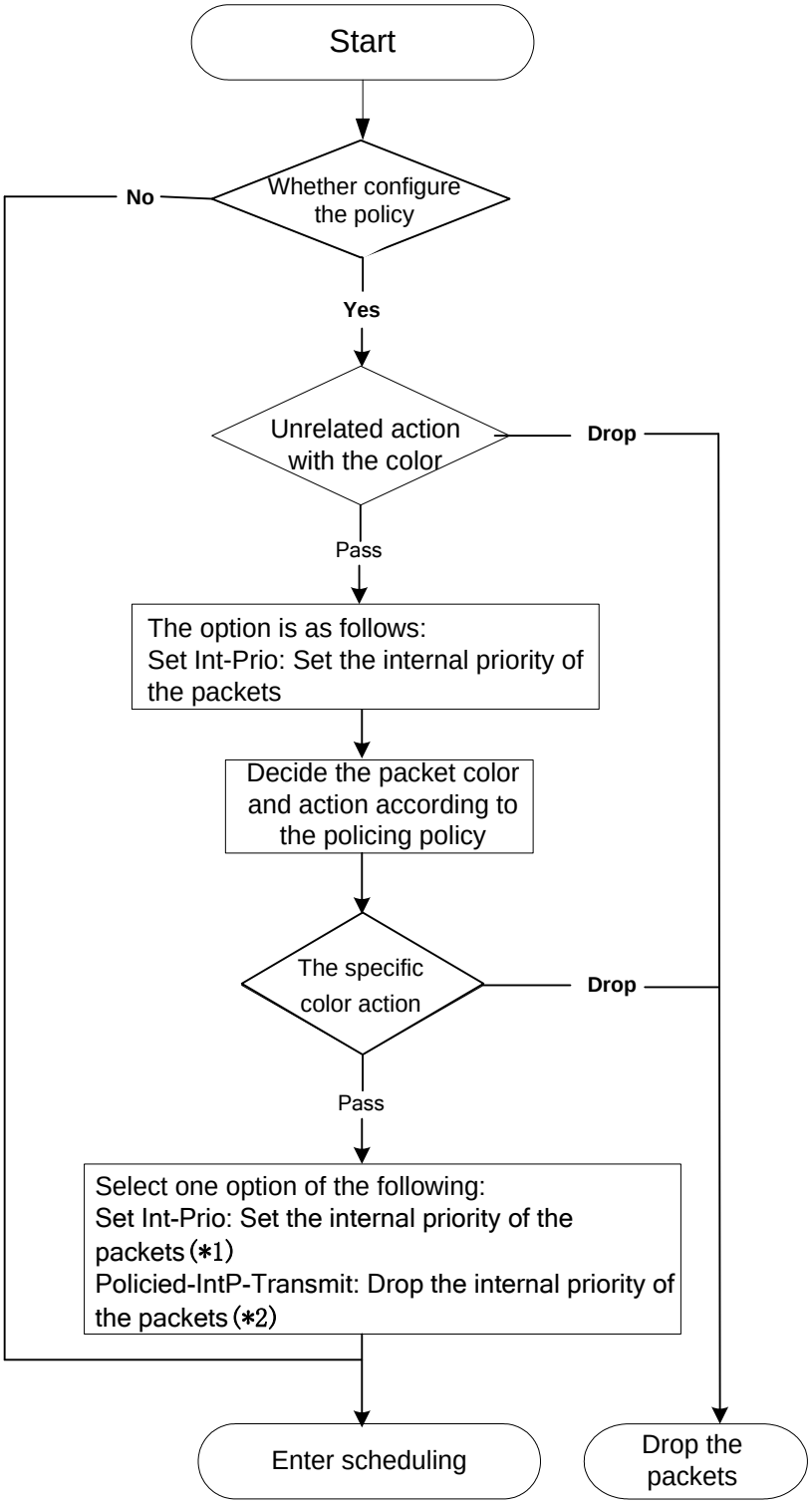
Classification process

Note 1: L2 CoS value is considered a property of the packets, there is no relation with the internal priority obtained of the following flow.

Note 2: Allow Trust DSCP and Trust COS to be configured at the same time, the priority is as follows: DSCP>COS.

Policing and remark: Each packet in classified ingress traffic is assigned an internal priority value, and can be policed and remarked.

Policing can be performed based on the flow to configure different policies that allocate bandwidth to classified traffic, the assigned bandwidth policy may be single bucket dual color or dual bucket three color. The traffic, will be assigned with different color, can be discarded or passed, for the passed packets, add the remarking action. Remarking uses a new Int-Prio value of lower priority to replace the original higher level Int-Prio value in the packet. COS and DSCP fields will be modifed according to the new Int-Prio at the egress. The following flowchart describes the operations.



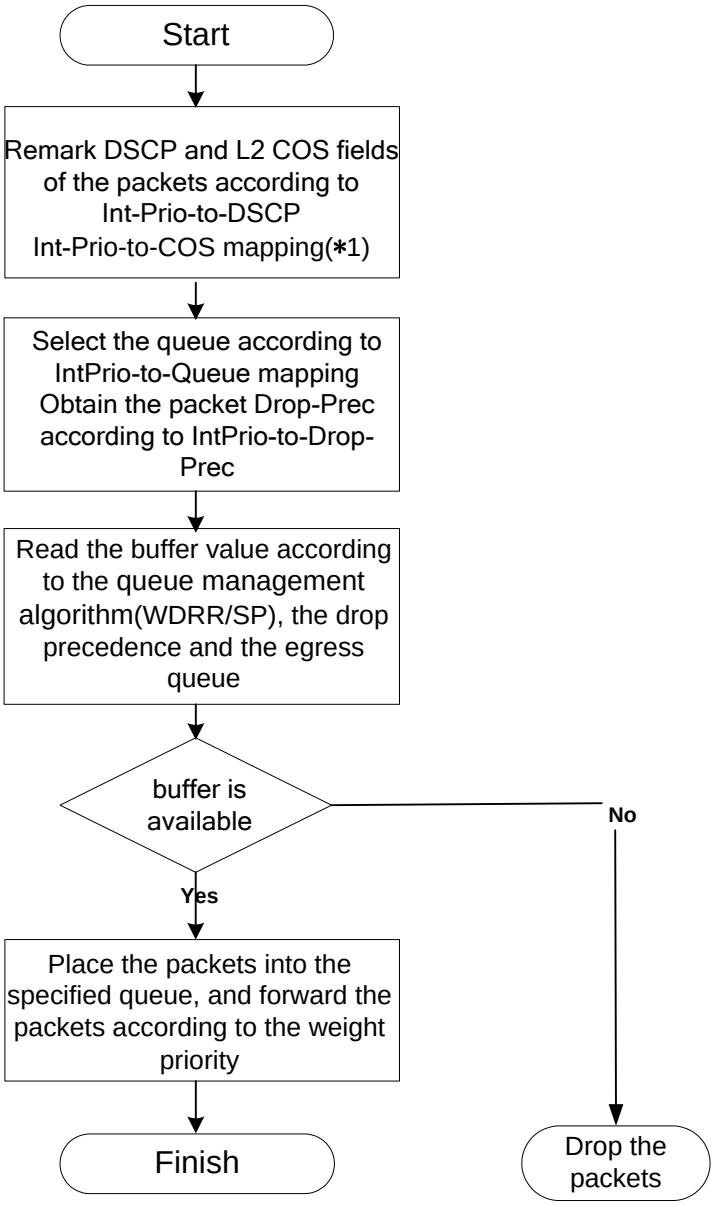
Policing and Remarking process

Note 1: Int-Prio will be covered with the after setting, Set Int-Prio of the specific color action will cover Set Int-Prio of the unrelated action with the color.

Note 2: Drop the internal priority of the packets according to IntP-to-IntP map. Source Int-Prio means to the obtainable Int-Prio in Classification flow or Int-Prio set by the unrelated action with the color.

Queuing and scheduling: There are the internal priority for the egress packets, the

scheduling operation assigns the packets to different priority queues according to the internal priority, and then forward the packets according to the priority queue weight and the drop precedence. The following flowchart describes the scheduling operation.



Queuing and Scheduling process

Note 1: The ingress configures pass-through-cos, pass-through-dscp to forbid the rewrite of L2 CoS priority and dscp value. At the egress, obtain L2 CoS priority and dscp value according to the final Int-Prio of the packets, decide whether rewrite L2 CoS priority and dscp value according to pass-through-cos, pass-through-dscp.

12.2 QoS Configuration Task List

Configure class map

Set up a classification rule according to ACL, CoS, VLAN ID, IPv4 Precedent, DSCP, IPV6 FL

to classify the data stream. Different classes of data streams will be processed with different policies.

Configure a policy map

After data steam classification, a policy map can be created to associate with the class map created earlier and enter class mode. Then different policies (such as bandwidth limit, priority degrading assigning new DSCP value) can be applied to different data streams. You can also define a policy set that can be use in a policy map by several classes.

Apply QoS to the ports or the VLAN interfaces

Configure the trust mode for ports or bind policies to ports. A policy will only take effect on a port when it is bound to that port.

The policy may be bound to the specific VLAN.

It is not recommended to synchronously use policy map on VLAN and its port, or else the policy map priority of the port is higher.

Configure queue management algorithm

Configure queue management algorithm, such as sp, wdrr, and so on.

Configure QoS mapping

Configure the mapping from CoS to DP, DSCP to DSCP, IntP or DP, IntP to DSCP.

1. Configure class map.

Command	Explanation
Global Mode	
class-map <class-map-name> **no class-map <class-map-name>**	Create a class map and enter class map mode; the "**no class-map <*class-map-name*>**" command deletes the specified class map.
match {access-group <acl-index-or-name> | ip dscp <dscp-list>| ip precedence <ip-precedence-list>| ipv6 access-group <acl-index-or-name> | ipv6 dscp <dscp-list>| ipv6 flowlabel <flowlabel-list>|vlan <vlan-list> | cos <*cos-list*>} **no match {access-group | ip dscp | ip precedence | ipv6 access-group | ipv6 dscp | ipv6 flowlabel | vlan | cos}**	Set matching criterion (classify data stream by ACL, CoS, VLAN ID, IPv4 Precedent, IPv6 FL or DSCP, etc) for the class map; the no command deletes specified matching criterion.

2. Configure a policy map

Command	Explanation
Global Mode	
policy-map <policy-map-name>	Create a policy map and enter policy map

no policy-map <policy-map-name>	mode; the no command deletes the specified policy map.
class <*class-map-name*> [insert-before <class-map-name>] **no class <class-map-name>**	After a policy map is created, it can be associated to a class. Different policy or new DSCP value can be applied to different data streams in class mode; the no command deletes the specified class.
set internal priority <new-intp> **no set internal priority**	Assign a new internal priority for the classified traffic; the no command cancels the new assigned value.
Single bucket mode: **policy <bits_per_second> <normal_burst_bytes> ({exceed-action ACTION})** **Dual bucket mode:** **policy <bits_per_second> <normal_burst_bytes> [pir <peak_rate_bps>] | <maximum_burst_bytes> [{exceed-action ACTION | violate-action ACTION }]** **ACTION definition:** **drop | transmit | set-internal-priority <intp_value> | policied-intp-transmit** **no policy**	Configure a policy for the classified flow. The non-aggregation policy command supports three colors. Analyze the working mode of the token bucket, whether it is singe rate single bucket, single rate dual bucket, dual rate dual bucket, set corresponding action to different color packets. The no command will delete the mode configuration.
policy aggregate <aggregate-policy-name> **no policy aggregate**	Apply a policy to classified traffic; the no command deletes the specified policy set.
accounting **no accounting**	Set statistic function for the classified traffic. After enable this function under the policy class map mode, add statistic function to the traffic of the policy class map. In single bucket mode, the messages can only red or green when passing policy. In the print information, there are two colors(green and red) of the packets. In dual bucket mode, there are three colors(green, red and yellow) of the packets.
Policy class map configuration mode	
drop **no drop**	Drop or transmit the traffic that match the class, the no command cancels the

transmit **no transmit**	assigned action.

3. Apply QoS to port or VLAN interface

Command	Explanation
Interface Configuration Mode	
mls qos trust {cos | dscp} **no mls qos trust {cos | dscp}**	Configure port trust; the no command disables the current trust status of the port.
mls qos cos {<*default-cos*>} **no mls qos cos**	Configure the default CoS value of the port; the no command restores the default setting.
service-policy input <*policy-map-name*> **no service-policy input <*policy-map-name*>**	Apply a policy map to the specified port; the no command deletes the specified policy map applied to the port. Egress policy map is not supported yet.
Global Mode	
service-policy input <policy-map-name> vlan <vlan-list> **no service-policy input <policy-map-name> vlan <vlan-list>**	Apply a policy map to the specified VLAN interface; the no command deletes the specified policy map applied to the VLAN interface.

4. Configure queue management algorithm and weight

Command	Explanation
Port Configuration Mode	
mls qos queue algorithm {sp | wdrr} **no mls qos queue algorithm**	Set queue management algorithm, the default queue management algorithm is wdrr.
Global Mode	
mls qos queue wdrr weight <weight0..weight7> **no mls qos queue wdrr weight**	Set wdrr queue weight for all ports globally, the default queue weight is 1 1 1 1 1 1 1 1.

5. Configure QoS mapping

Command	Explanation
Global Mode	
mls qos map (cos-dp <dp1…dp8> | dscp-dscp	Set the priority mapping for QoS, the no

<in-dscp list> to <out-dscp> | dscp-intp <in-dscp list> to <intp> | dscp-dp <in-dscp list> to <dp>) **no mls qos map (cos-dp | dscp-dscp | dscp-intp | dscp-dp)** **mls qos map intp-dscp <dscp1..dscp8>** **no mls qos map intp-dscp**	command restores the default mapping value.

6. Clear accounting data of the specific ports or VLANs

Command	Explanation
Admin Mode	
clear mls qos statistics [interface <interface-name> | vlan <vlan-id>]	Clear accounting data of the specified ports or VLAN Policy Map. If there are no parameters, clear accounting data of all policy map.

7. Show configuration of QoS

Command	Explanation
Admin Mode	
show mls qos maps [cos-dp | dscp-dscp | dscp-intp | dscp-dp | intp-dscp]	Display the configuration of QoS mapping.
show class-map [<class-map-name>]	Display the classified map information of QoS.
show policy-map [<policy-map-name>]	Display the policy map information of QoS.
show mls qos {interface [<interface-id>] [policy | queuing] | vlan <vlan-id>}	Displays QoS configuration information on a port.

12.3 QoS Example

Example 1:

Enable QoS function, change the queue out weight of port ethernet 1/0/1 to 1:1:2:2:4:4:8:8, set the port in trust CoS mode without changing DSCP value, and set the default CoS value of the port to 5.

The configuration steps are listed below:

Switch#config

Switch(config)# mls qos queue weight 1 1 2 2 4 4 8 8

Switch(Config-If-Ethernet 1/0/1)#mls qos trust cos

Switch(Config-If-Ethernet1/0/1)#mls qos cos 5

Configuration result:

When QoS enabled in Global Mode, the egress queue bandwidth proportion of each port is 1:1:2:2:4:4:8:8. When packets have CoS value coming in through port ethernet1/0/1, it will be map to the internal priority according to the CoS value, CoS value 0 to 7 correspond to queue out 1, 2, 3, 4, 5, 6, 7, 8 respectively. If the incoming packet has no CoS value, it is default to 5 and will be put in queue6. All passing packets would not have their DSCP values changed

Example 2:

In port ethernet1/0/2, set the bandwidth for packets from segment 192.168.1.0 to 10 Mb/s, with a burst value of 4 MB, all packets exceed this bandwidth setting will be dropped.

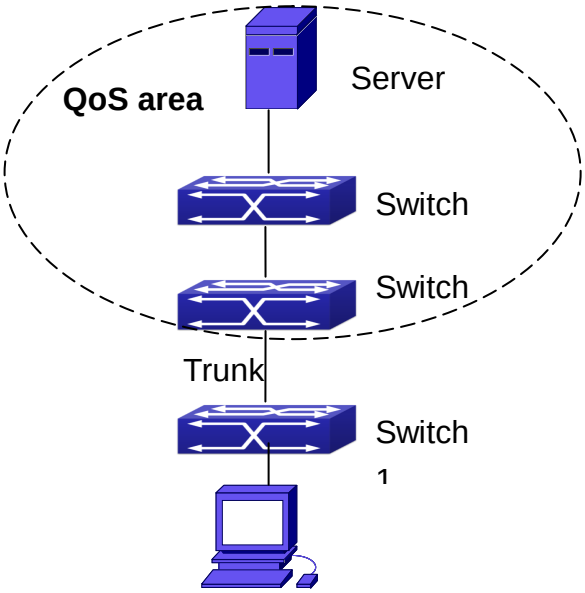
The configuration steps are listed below:

Switch#config
Switch(config)#access-list 1 permit 192.168.1.0 0.0.0.255
Switch(config)#class-map c1
Switch(Config-ClassMap-c1)#match access-group 1
Switch(Config-ClassMap-c1)#exit
Switch(config)#policy-map p1
Switch(Config-PolicyMap-p1)#class c1
Switch(Config-PolicyMap-p1-Class-c1)#policy 10000 4000 exceed-action drop
Switch(Config-PolicyMap-p1-Class-c1)#exit
Switch(Config-PolicyMap-p1)#exit
Switch(config)#interface ethernet 1/0/2
Switch(Config-If-Ethernet1/0/2)#service-policy input p1

Configuration result:

An ACL name 1 is set to matching segment 192.168.1.0. Enable QoS globally, create a class map named c1, matching ACL1 in class map; create another policy map named p1 and refer to c1 in p1, set appropriate policies to limit bandwidth and burst value. Apply this policy map on port ethernet1/0/2. After the above settings done, bandwidth for packets from segment 192.168.1.0 through port ethernet 1/0/2 is set to 10 Mb/s, with a burst value of 4 MB, all packets exceed this bandwidth setting in that segment will be dropped.

Example 3:



Typical QoS topology

As shown in the figure, inside the block is a QoS domain, Switch1 classifies different traffics and assigns different IP precedences. For example, set CoS precedence for packets from segment 192.168.1.0 to 5 on port ethernet1/0/1(set the internal priority to 40, set the default intp-dscp mapping to 40-40, the corresponding IP precedence to 5). The port connecting to switch2 is a trunk port. In Switch2, set port ethernet 1/0/1 that connecting to swtich1 to trust dscp. Thus inside the QoS domain, packets of different priorities will go to different queues and get different bandwidth.

The configuration steps are listed below:
QoS configuration in Switch1:
Switch#config
Switch(config)#access-list 1 permit 192.168.1.0 0.0.0.255
Switch(config)#class-map c1
Switch(Config-ClassMap-c1)#match access-group 1
Switch(Config-ClassMap-c1)#exit
Switch(config)#policy-map p1
Switch(Config-PolicyMap-p1)#class c1
Switch(Config-PolicyMap-p1-Class-c1)#set ip precedence 40
Switch(Config-PolicyMap-p1-Class-c1)#exit
Switch(Config-PolicyMap-p1)#exit
Switch(config)#interface ethernet 1/0/1
Switch(Config-If-Ethernet1/0/1)#service-policy input p1

QoS configuration in Switch2:
Switch#config

Switch(config)#interface ethernet 1/0/1
Switch(Config-If-Ethernet1/0/1)#mls qos trust cos

12.4 QoS Troubleshooting

trust cos and EXP can be used with other trust or Policy Map.
trust dscp can be used with other trust or Policy Map. This configuration takes effect to IPv4 and IPv6 packets.
trust exp, trust dscp and trust cos may be configured at the same time, the priority is: EXP>DSCP>COS.
If the dynamic VLAN (mac vlan/voice vlan/ip subnet vlan/protocol vlan) is configured, then the packet COS value equals COS value of the dynamic VLAN.
At present, it is not recommended to synchronously use policy map on VLAN and VLAN's port.

Chapter 13 Flow-based Redirection

13.1 Introduction to Flow-based Redirection

Flow-based redirection function enables the switch to transmit the data frames meeting some special condition (specified by ACL) to another specified port. The fames meeting a same special condition are called a class of flow, the ingress port of the data frame is called the source port of redirection, and the specified egress port is called the destination port of redirection. Usually there are two kinds of application of flow-based redirection: 1. connecting a protocol analyzer (for example, Sniffer) or a RMON monitor to the destination port of redirection, to monitor and manage the network, and diagnose the problems in the network; 2. Special transmission policy for a special type of data frames.

The switch can only designate a single destination port of redirection for a same class of flow within a source port of redirection, while it can designate different destination ports of redirection for different classes of flows within a source port of redirection. The same class of flow can be applied to different source ports.

13.2 Flow-based Redirection Configuration Task Sequence

Flow-based redirection configuration

Check the current flow-based redirection configuration

1. Flow-based redirection configuration

Command	Explanation
Physical Interface Configuration Mode	
access-group <*aclname*> redirect to interface [ethernet <*IFNAME*>|<*IFNAME*>] **no access-group <*aclname*> redirect**	Specify flow-based redirection for the port; the "**no access-group <*aclname*> redirect**" command is used to delete flow-based redirection.

2. Check the current flow-based redirection configuration

Command	Explanation
Global Mode/Admin Mode	
show flow-based-redirect {interface [ethernet <*IFNAME*> |<*IFNAME*>]}	Display the information of current flow-based redirection in the system/port.

13.3 Flow-based Redirection Examples

Example:
User's request of configuration is listed as follows: redirecting the frames whose source IP is 192.168.1.111 received from port 1 to port 6, that is sending the frames whose source IP is 192.168.1.111 received from port 1 through port6.

Modification of configuration:
1: Set an ACL, the condition to be matched is: source IP is 192.168.1.111;
2: Apply the redirection based on this flow to port 1.

The following is the configuration procedure:
Switch(config)#access-list 1 permit host 192.168.1.111
Switch(config)#interface ethernet 1/0/1
Switch(Config-If-Ethernet1/0/1)# access-group 1 redirect to interface ethernet 1/0/6

13.4 Flow-based Redirection Troubleshooting Help

When the configuration of flow-based redirection fails, please check that whether it is the following reasons causing the problem:
The type of flow (ACL) can only be digital standard IP ACL, digital extensive IP ACL, nomenclature standard IP ACL, nomenclature extensive IP ACL, digital standard IPv6 ACL, and nomenclature standard IPv6 ACL;
Parameters of Timerange and Portrange can not be set in ACL, the type of ACL should be Permit.
The redirection port must be 1000Mb port in the flow-based redirection function.
Do not implement the forward across VLAN for flow-based redirection.

Chapter 14 Egress QoS Configuration

14.1 Introduction to Egress QoS

In traditional IP networks, all packets are treated in the same way. All network equipments treat them by the first-in-first-out policy and try best effort to send them to the destination. However, it does not guarantee the performance like reliability and transmission delay. Network develops so fast that new demand has been raised for the quality of service on IP network with the continual emergence of new applications. For example, delay-sensitive services like VoIP and video put higher demands on packet transmission delay and users cannot accept too long transmission delay (by contrast, E-mail and FTP services are not sensitive to the time delay). In order to support services with different service requirement like voice, video and data service, the network is required to be able to distinguish between different communications and provide appropriate service. The traditional best-effort IP network cannot identify and distinguish various kinds of communications while this ability is the very premise of providing differentiated services for different communications. Therefore, the best-effort service mode of traditional network cannot meet the demand of applications. The emergence of QoS techniques is committed to solve this problem.

Egress PolicyMap is the QoS policy in egress which performs QoS control of packets in the egress direction and provides better service for specified network communication with kinds of techniques. Egress PolicyMap includes class-map and policy-map, of which class-map is used for selecting packets to operate and policy-map is used for specifying the operation to use. Not all equipments support Egress QoS currently.

14.1.1 Egress QOS Terms

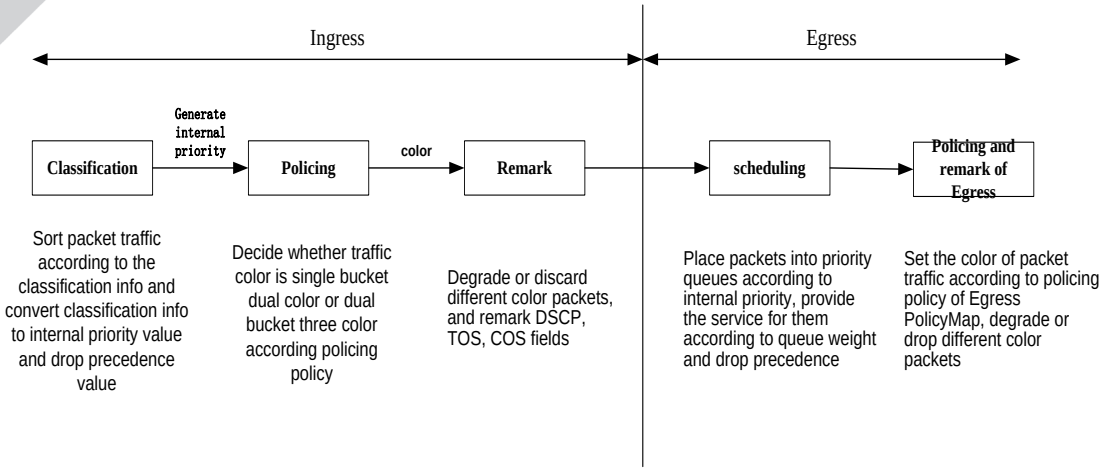
Egress QoS: Achieving QoS on egress of port.

Inner_vid: VLAN ID brought by the TAG near the header of network layer when double TAGs exist.

Outer_vid: VLAN ID brought by the TAG near the header of network link layer when double TAGs exist. The TAG is considered to be outer tag by default when only one TAG exists.

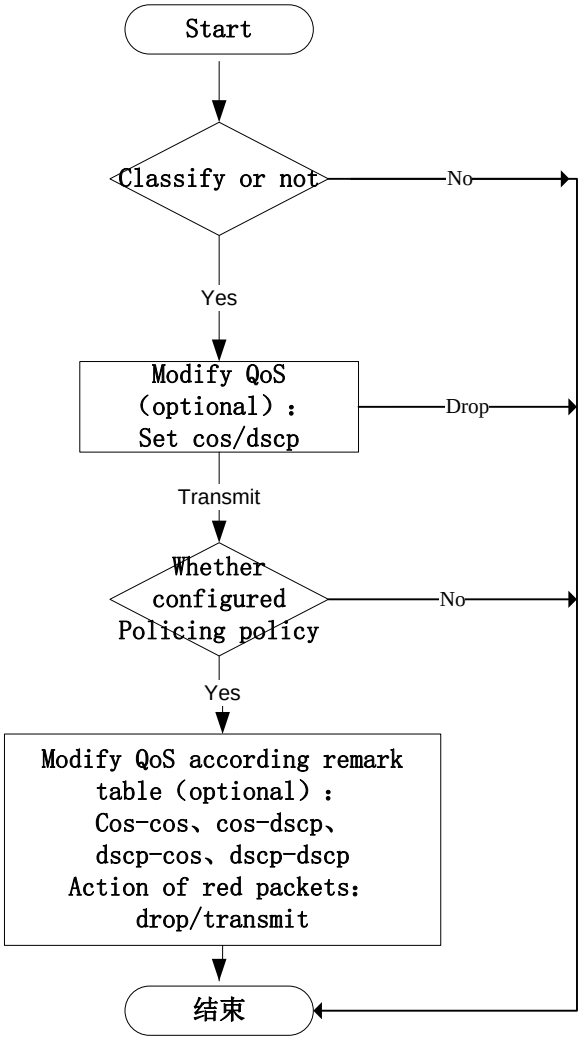
Outer_tpid: Protocol type of the network link layer header indicating the type of outer tag.

14.1.2 Basic Egress QoS Model



According to the characters (including field values like COS and DSCP) of upstream packets, policing and rewriting of Egress make the last QoS change on the packet prior to the packet egress.

Policing configures different policing policy based on the flow and distributes bandwidth for the flow classified. The distribution policy of bandwidth can be either dual bucket dual color or dual bucket three color. Different colors can be assigned to different flows and approaches of discard or passage can be chosen for them; you can add rewriting action for packets with passage approach chosen. See the following flow chart for detailed description of Egress QoS:



Description of action that modify QoS attribute according to egress remark table:

cos-cos: for cos value of packets, modify cos value of packets according to cos table of QoS remarking

cos-dscp: for cos value of packets, modify dscp value of packets according to cos table of QoS remarking

dscp-cos: for dscp value of packets, modify cos value of packets according to dscp table of QoS remarking

dscp-dscp: for dscp value of packets, modify dscp value of packets according to dscp table of QoS remarking

14.2 Egress QoS Configuration

Egress QoS Configuration Task List:

Configure class map

Set up a classification rule according to ACL, CoS, VLAN ID, IPv4 Precedent, DSCP, IPV6 DSCP to classify the data stream. Different classes of data streams will be processed with different policies.

Configure policy map

After data steam classification, a policy map can be created to associate with a class map created earlier and enter policy class mode. Then different policies (such as bandwidth limit, assigning new DSCP value) can be applied to different data streams.

Apply Egress QoS to port or VLAN

Configure the trust mode or binding policies for ports. A policy will only take effect on a port when it is bound to that port.

The policy may be bound to the specific VLAN.

Set Egress QoS remark mapping

If modify QoS attribute by using Egress QoS remark in policy, it should set the corresponding mapping. If it needs to take effect to green packets, modifying switch of green packets should be enabled and ingress needs to trust the corresponding QoS attribute (qos/dscp/exp).

1. Configure a class-map

Command	Explanation
Global Mode	
class-map <class-map-name> **no class-map <class-map-name>**	Create a class-map and enter class-map mode, no command deletes the specified class-map.
match {access-group <*acl-index-or-name*> | ip dscp <*dscp-list*> | ip precedence <*ip-precedence-list*> | ipv6 dscp <*dscp-list*> | vlan <*vlan-list*> | cos <*cos-list*> | ipv6 access-group <*acl-index-or-name*>} **no match {access-group | ip dscp | ip precedence | ipv6 dscp | vlan | cos | ipv6 access-group}**	Configure the matched standard of the class map to classify the data stream according to ACL, CoS, VLAN ID, IPv4 Precedence, DSCP, IPv6 DSCP priority; no command deletes the specific matched standard.

2. Configure a policy-map

Command	Explanation
Global Mode	
policy-map <policy-map-name> **no policy-map <policy-map-name>**	Create a policy-map and enter policy-map mode, no command deletes the specific policy-map.
class <class-map-name> [insert-before <class-map-name>] **no class <class-map-name>**	Create a policy map to associate with a class map and enter policy class map mode, then different data streams can apply different policies and be assigned a new DSCP value. No command deletes the specified policy class map.

set {ip dscp <*new-dscp*> | ip precedence <*new-precedence*> | cos <*new-cos*> | c-vid <*new-c-vid*> | s-vid <*new-s-vid*> | s-tpid <*new-s-tpid*>} **no set {ip dscp | ip precedence | cos | c-vid | s-vid | s-tpid}**	Assign a new DSCP, CoS and IP Precedence value for the classified flow, no command cancels the operation.
Single bucket mode: **policy <*bits_per_second*> <*normal_burst_bytes*> ({action ACTION} | exceed-action drop | transmit})** **Dual bucket mode:** **policy <*bits_per_second*> <*normal_burst_bytes*> [pir <*peak_rate_bps*>] | <*maximum_burst_bytes*> [{action ACTION | violate-action drop | transmit}]** **ACTION definition:** **policied-cos-to-cos-transmit | policied-cos-to-dscp-transmit | policied-dscp-exp-to-cos-transmit | policied-dscp-exp-to-dscp-transmit** **no policy**	Configure a policy for the classified flow. The non-aggregation policy command supports three colors. Analyze the working mode of the token bucket, whether it is single rate single bucket, single rate dual bucket or dual rate dual bucket, set corresponding action to different color packets. The no command will delete the configuration. Only specific switch supports single bucket mode.
accounting **no accounting**	Set statistic function for the classified flow. After enable this function under the policy class map mode, add statistic function to the flow of the policy class map. In single bucket mode, packets can only red or green when passing policy. In the print information, in-profile means green and out-profile means red. In dual bucket mode, there are three colors of packets in-profile means green and out-profile means red and yellow.

3. Apply policy to port or VLAN

Command	Explanation
Interface Mode	
service-policy output <*policy-map-name*>	Apply a policy map to the egress of the port;

no service-policy output<*policy-map-name*>	the no command deletes the specified policy map applied to the port.
Global Mode	
service-policy output <policy-map-name> vlan <vlan-list> **no service-policy output <*policy-map-name*> vlan <*vlan-list*>**	Apply a policy map to the egress of the VLAN; the no command deletes the specified policy map applied to the VLAN interface.

4. Set Egress QoS remark mapping

Command	Explanation
Global Mode	
mls qos map {cos-cos | cos-dscp} {green | yellow | red} <value1> <value2>…<value8> **no mls qos map {cos-cos | cos-dscp} {green | yellow | red}**	Set Egress cos mapping, no command resotores the default configuration.
mls qos map {dscp-cos | dscp-dscp} {green | yellow | red} <dscp list> to <value> **no mls qos map {dscp-cos | dscp-dscp} {green | yellow | red}**	Set Egress dscp mapping, <dscp-list> means 1 to 8 dscp values, no command restores the default configuration.
mls qos egress green remark **no mls qos egress green remark**	Set Egress QoS remark mapping to take effect for green packets, no command does not take effect to green packets.

5. Clear accounting data of the specific ports or VLANs

Command	Explanation
Admin Mode	
clear mls qos statistics [interface <interface-name> | vlan <vlan-id>]	Clear accounting data of the specified ports or VLAN Policy Map. If there are no parameters, clear accounting data of all policy map.

6. Show QoS configuration

Command	Explanation
Admin Mode	
show mls qos {interface [<*interface-id*>] [policy | queuing] | vlan <*vlan-id*>}	Show QoS configuration of the port.
show class-map [<*class-map-name*>]	Show the class map information of QoS.
show policy-map [<*policy-map-name*>]	Show the policy map information of QoS.

show mls qos maps {cos-cos | cos-dscp | dscp-cos | dscp-exp} {green | yellow | red |}	Show mapping relation of Egress QoS remark.

14.3 Egress QoS Examples

Example1:
On the egress of the port1, change cos value as 4 for the packet with dscp value of 0.
Create a class map:
switch(config)#class-map 1
switch(config-classmap-1)#match ip dscp 0
switch(config-classmap-1)#exit

Create a policy map:
switch(config)#policy-map 1
switch(config-policymap-1)#class 1
switch(config-policymap-1-class-1)#set cos 4
switch(config-policymap-1-class-1)#exit
switch(config-policymap-1)#exit

Bind a policy to the port:
switch(config)#in e 1/0/1
switch(config-if-ethernet1/0/1)#service-policy output 1

Example2:
On the egress of vlan10, change cos value as 4 for the packet with ipv6 dscp value of 7.
Create a class map:
switch(config)#class-map 1
switch(config-classmap-1)#match ipv6 dscp 7
switch(config-classmap-1)#exit

Create a policy map:
switch(config)#policy-map 1
switch(config-policymap-1)#class 1
switch(config-policymap-1-class-1)#set cos 4
switch(config-policymap-1-class-1)#exit
switch(config-policymap-1)#exit

Bind a policy to VLAN
switch(config)#service-policy output 1 vlan 10

Example 3:
In egress of port 1, limit the speed of packets. Set the bandwidth for packets to 1 Mb/s, with the normal burst value of 1 MB, the max burst value of 4 MB, set dscp value of 1 as 10 for green packets, set dscp value of yellow packets as 9 and drop red packets.
Create a class map
switch(config)#class-map c1
switch(config-classmap-c1)#match ip dscp 1
switch(config-classmap-c1)#exit
Create a policy map
switch(config)#policy-map p1
switch(config-policymap-p1)#class c1
switch(config-policymap-p1-class-c1)#policy 1000 1000 4000 action policied-dscp-exp-to-dscp-transmit violate-action drop
switch(config-policymap-p1-class-c1)#exit
switch(config-policymap-p1)#exit
Set Egress dscp remark mapping
switch(config)#mls qos map dscp-dscp green 1 to 10
switch(config)#mls qos map dscp-dscp yellow 1 to 9
Set Egress remark to take effect for green packets
switch(config)#mls qos egress green remark
Set trust dscp mode on ingress
switch(config-if-port-range)#mls qos trust dscp
Bind policy to egress of port1
switch(config-if-ethernet1/0/1)#service-policy output p1

14.4 Egress QoS Troubleshooting Help

Not all equipments support Egress QoS presently, so please make sure the current device supports this function.
If the policy configured cannot bind to the port or VLAN, please check whether the match option in classification table is supported by the current device.
If terminal printing suggests lack of resource, please make sure there is enough resource to send the current policy.
If the policy with match acl configured cannot bind to the port or VLAN, please make sure rules including permit exist in ACL.
If modifying QoS attribute is invalid by Egress QoS remark, please ensure whether ingress sets the correspondsing QoS attribute with trust.
If egress set QoS attributes (set cos/ip dscp) for modifying all packets, and it uses Egress remark to modify QoS attributes for packets of different colors, previous modification is

preferential for modifying packets.

Chapter 15 Flexible QinQ Configuration

15.1 Introduction to Flexible QinQ

15.1.1 QinQ Technique

Dot1q-tunnel is also called QinQ (802.1Q-in-802.1Q), which is an expansion of 802.1Q. Its dominating idea is encapsulating the customer VLAN tag (CVLAN tag) to the service provider VLAN tag (SPVLAN tag). The packet with two VLAN tags is transmitted through the backbone network of the ISP internet to provide a simple layer-2 tunnel for the users. It is simple and easy to manage, applicable only by static configuration, and especially adaptive to small office network or small metropolitan area network using layer-3 switch as backbone equipment.
There are two kinds of QinQ: basic QinQ and flexible QinQ, the priority of flexible QinQ is higher than basic QinQ.

15.1.2 Basic QinQ

Basic QinQ based the port. After a port configures QinQ, whether the received packet with tag or not, the device still packs the default VLAN tag for the packet. Using basic QinQ is simple, but the setting method of VLAN tag is inflexible.

15.1.3 Flexible QinQ

Flexible QinQ based data flow. It selects whether pack the external tag and packs what kind of the external tag by matching the material flow. For example: implement the property of flexible QinQ according to the user's VLAN tag, MAC address, IPv4/IPv6 address, IPv4/IPv6 protocol and the port ID of the application, etc. So, it can encapsulate the external tag for the packet and implements different scheme by different users or methods.

15.2 Flexible QinQ Configuration Task List

The match of flexible QinQ data flow uses policy-map rule of QoS to be sent, the configuration task list is as follows:
1. Create class-map to classify different data flows
2. Create flexible QinQ policy-map to relate with the class-map and set the corresponding operation
3. Bind flexible QinQ policy-map to port

1. Configure class map

Command	Explanation
Global mode	

class-map <class-map-name> **no class-map <class-map-name>**	Create a class-map and enter class-map mode, the no command deletes the specified class-map.
match {access-group <*acl-index-or-name*> | ip dscp <*dscp-list*>| ip precedence <*ip-precedence-list*>| ipv6 access-group <*acl-index-or-name*>| ipv6 dscp <*dscp-list*> | ipv6 flowlabel <flowlabel-list> | vlan <vlan-list> | cos <*cos-list*>} **no match {access-group | ip dscp | ip precedence|ipv6 access-group| ipv6 dscp | ipv6 flowlabel | vlan | cos}**	Set the match standard of class-map, (classify data flow by ACL, CoS, VLAN ID, IPv4 Precedent or DSCP, etc for the class map); the no command deletes the specified match standard.

2. Configure policy-map of flexible QinQ

Command	Explanation
Global mode	
policy-map <policy-map-name> **no policy-map <policy-map-name>**	Create a policy-map and enter policy-map mode, the no command deletes the specified policy-map.
class <*class-map-name*> [insert-before <class-map-name>] **no class <class-map-name>**	After a policy-map is created, it can be associated to a class. Different policy or new DSCP value can be applied to different data flows in class mode; the no command deletes the specified class-map.
set s-vid < vid> **no set s-vid**	Set external VLAN Tag for the classified traffic, no command cancels the operation.
add s-vid <vid> **no add s-vid <vid>**	Add external VLAN Tag for the classified traffic, no command cancels the operation.

3. Bind flexible QinQ policy-map to port

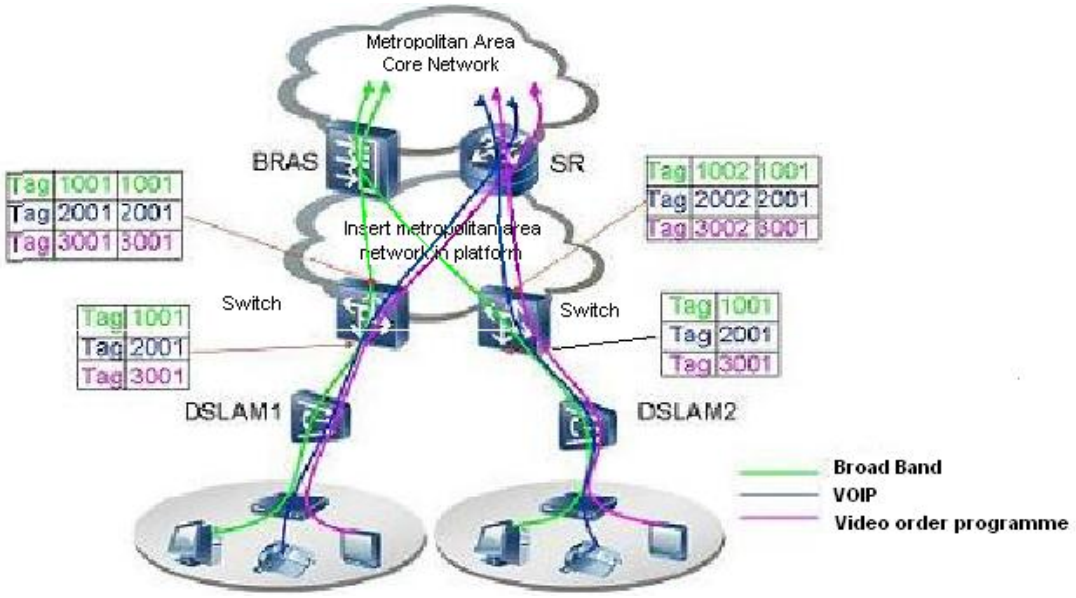
Command	Explanation
Port mode	
service-policy input<*policy-map-name*> **no service-policy input<*policy-map-name*>**	Apply a policy-map to a port, the no command deletes the specified policy-map applied to the port.

Global mode	
service-policy input<*policy-map-name*> vlan<vid> **no service-policy input<policy-map-name> vlan <vid>**	Apply a policy-map to a VLAN, the no command deletes the specified policy-map applied to the VLAN.

4. Show flexible QinQ policy-map bound to port

Command	Explanation
Admin mode	
show mls qos {interface [<interface-id>]	Show flexible QinQ configuration on the port.

15.3 Flexible QinQ Example



Flexible QinQ application topology

As shown in the figure, the first user is assigned three VLANs that the tag values are 1001, 2001, 3001 respectively in DSLAM1. VLAN1001 corresponds to Broad Band Network, VLAN2001 corresponds to VOIP, VLAN3001 corresponds to VOD. After the downlink port enables flexible QinQ function, the packets will be packed with different external tags according to VLAN ID of users. The packet with tag 1001 will be packed an external tag 1001 directly(This tag is unique in public network), enter Broad Band Network-VLAN1001 and classfied to BRAS device. The packet with tag 2001(or 3001) will be packed an external tag 2001(or 3001) and classfied to SR device according to the flow rules. The second user can be assigned different VLAN tags for different VLANs in DSLAM2. Notice: The assigned VLAN tag of the second user may be same with the first user and the packet with tag will be also packed

an external tag. In the above figure, the external tag of the second user is different to the first user for distinguishing DSLAM location and locating the user finally.

The configuration in the following:

If the data flow of DSLAM1 enters the switch's downlink port1, the configuration is as follows:

```
Switch(config)#class-map c1
Switch(config-classmap-c1)#match vlan 1001
Switch(config-classmap-c1)#exit
Switch(config)#class-map c2
Switch(config-classmap-c2)#match vlan 2001
Switch(config-classmap-c2)#exit
Switch(config)#class-map c3
Switch(config-classmap-c3)#match vlan 3001
Switch(config-classmap-c3)#exit
Switch(config)#policy-map p1
Switch(config-policymap-p1)#class c1
Switch(config-policymap-p1-class-c1)# set s-vid 1001
Switch(config-policymap-p1)#class c2
Switch(config-policymap-p1-class-c2)# set s-vid 2001
Switch(config-policymap-p1)#class c3
Switch(config-policymap-p1-class-c3)# set s-vid 3001
Switch(config-policymap-p1-class-c3)#exit
Switch(config-policymap-p1)#exit
Switch(config)#interface ethernet 1/0/1
Switch(config-if-ethernet1/0/1)#service-policy input p1
```

If the data flow of DSLAM2 enters the switch's downlink port1, the configuration is as follows:

```
Switch(config)#class-map c1
Switch(config-classmap-c1)#match vlan 1001
Switch(config-classmap-c1)#exit
Switch(config)#class-map c2
Switch(config-classmap-c2)#match vlan 2001
Switch(config-classmap-c2)#exit
Switch(config)#class-map c3
Switch(config-classmap-c3)#match vlan 3001
Switch(config-classmap-c3)#exit
Switch(config)#policy-map p1
Switch(config-policymap-p1)#class c1
Switch(config-policymap-p1-class-c1)# set s-vid 1002
Switch(config-policymap-p1)#class c2
Switch(config-policymap-p1-class-c2)# set s-vid 2002
```

Switch(config-policymap-p1)#class c3
Switch(config-policymap-p1-class-c3)# set s-vid 3002
Switch(config-policymap-p1-class-c3)#exit
Switch(config-policymap-p1)#exit
Switch(config)#interface ethernet 1/0/1
Switch(config-if-ethernet1/0/1)# service-policy input p1

15.4 Flexible QinQ Troubleshooting

If flexible QinQ policy can not be bound to the port, please check whether the problem is caused by the following reasons:
Make sure flexible QinQ whether supports the configured class-map and policy-map
Make sure ACL includes permit rule if the class-map matches ACL rule
Make sure the switch exists enough TCAM resource to send the binding
Priority of flexible QinQ and vlan ingress filtering for processing packets is: flexible QinQ > vlan ingress filtering

Chapter 16 Layer 3 Forward Configuration

Switch supports Layer 3 forwarding which forwards Layer 3 protocol packets (IP packets) across VLANs. Such forwarding uses IP addresses, when a interface receives an IP packet, it will perform a lookup in its own routing table and decide the operation according to the lookup result. If the IP packet is destined to another subnet reachable from this switch, then the packet will be forwarded to the appropriate interface. Switch can forward IP packets by hardware, the forwarding chip of switch have a host route table and default route table. Host route table stores host routes to connect to the switch directly; default route table stores network routes (after aggregation algorithm process).

If the route (either host route or network route) for forwarding unicast traffic exists in the forwarding chip, the forwarding of traffic will be completely handled by hardware. As a result, forwarding efficiency can be greatly improved, even to wire speed.

16.1 Layer 3 Interface

16.1.1 Introduction to Layer 3 Interface

Layer 3 interface can be created on switch. The Layer 3 interface is not a physical interface but a virtual interface. Layer 3 interface is built on VLANs. The Layer 3 interface can contain one or more layer 2 ports which belong to the same VLAN, or contain no layer 2 ports. At least one of the Layer 2 ports contained in Layer 3 interface should be in UP state for Layer 3 interface in UP state, otherwise, Layer 3 interface will be in DOWN state. All layer 3 interfaces in the switch use the same MAC address by default, this address is selected from the reserved MAC address while creating Layer 3 interface. The Layer 3 interface is the base for layer 3 protocols. The switch can use the IP addresses set in the layer 3 interfaces to communicate with the other devices via IP. The switch can forward IP packets between different Layer 3 interfaces. Loopback interface belongs to Layer 3 interface.

16.1.2 Layer 3 Interface Configuration Task List

Layer 3 Interface Configuration Task List:

1. Create Layer 3 interface
2. Bandwidth for Layer 3 Interface configuration
3. Configure VLAN interface description
4. Open or close the VLAN interface
5. VRF configuration

(1) Create VRF instance and enter VPN view

(2) Configure RD of VRF instance (optional)

(3) Configure RT of VRF instance (optional)

(4) Configure the relation between VRF instance and the interface

1. Create Layer 3 Interface

Command	Explanation
Global Mode	
interface vlan <*vlan-id*> **no interface vlan <*vlan-id*>**	Creates a VLAN interface (VLAN interface is a Layer 3 interface); the no command deletes the VLAN interface (Layer 3 interface) created in the switch.
interface loopback <*loopback-id*> **no interface loopback <*loopback-id*>**	Creates a Loopback interface then enter the loopback Port Mode; the no command deletes the Loopback interface created in the switch.

2. Bandwidth for Layer 3 Interface configuration

Command	Explanation
VLAN Interface Mode	
bandwidth <bandwidth> **no bandwidth**	Configure the bandwidth for Layer 3 Interface. The no command recovery the default value.

3. Configure VLAN interface description

Command	Explanation
VLAN Interface Mode	
description <*text*> **no description**	Configure the description information of VLAN interface. The no command will cancel the description information of VLAN interface.

4. Open or close the vlan interface

Command	Explanation
VLAN Interface Mode	
shutdown **no shutdown**	Open or close the vlan interface.

5. VRF configuration

(1) Create VRF instance and enter VPN view

(2) Configure RD of VRF instance (optional)

(3) Configure RT of VRF instance (optional)

(4) Configure the relation between VRF instance and the interface

Command	Explanation
Global Mode	
ip vrf <vrf-name> **no ip vrf <vrf-name>**	Create VRF instance; VRF instance is not created by default.
VRF Mode	
rd <ASN:nn_or_IP-address:nn>	Configure RD of VRF instance. RD is not created by default.
route-target {import | export | both} <rt-value> **no route-target {import | export | both} <rt-value>**	Configure RT of VRF instance
Interface Mode	
ip vrf forwarding <vrf-name> **no ip vrf forwarding <vrf-name>**	Configure the relation between VRF instance and the interface.
ip address <ip-address> <mask> **no ip address <ip-address> <mask>**	Configure the private IP address of direct link interface.

16.2 IP Configuration

16.2.1 Introduction to IPv4, IPv6

IPv4 is the current version of global universal Internet protocol. The practice has proved that IPv4 is simple, flexible, open, stable, strong and easy to implement while collaborating well with various protocols of upper and lower layers. Although IPv4 almost has not been changed since it was established in 1980's, it has kept growing to the current global scale with the promotion of Internet. However, as Internet infrastructure and Internet application services continue boosting, IPv4 has shown its deficiency when facing the present scale and complexity of Internet.

IPv6 refers to the sixth version of Internet protocol which is the next generation Internet protocol designed by IETF to replace the current Internet protocol version 4 (IPv4). IPv6 was specially developed to make up the shortages of IPv4 addresses so that Internet can develop further.

The most important problem IPv6 has solved is to add the amount of IP addresses. IPv4 addresses have nearly run out, whereas the amount of Internet users has been increasing in geometric series. With the greatly and continuously boosting of Internet services and application devices (Home and Small Office Network, IP phone and Wireless Service Information Terminal which make use of Internet,) which require IP addresses, the supply of IP addresses turns out to be more and more tense. People have been working on the problem of shortage of IPv4 addresses for a long time by introducing various technologies to prolong the

lifespan of existing IPv4 infrastructure, including Network Address Translation(NAT for short), and Classless Inter-Domain Routing(CIDR for short), etc.

Although the combination of CIDR, NAT and private addressing has temporarily mitigated the problem of IPv4 address space shortage, NAT technology has disrupted the end-to-end model which is the original intention of IP design by making it necessary for router devices that serve as network intermediate nodes to maintain every connection status which increases network delay greatly and decreases network performance. Moreover, the translation of network data packet addresses baffles the end-to-end network security check, IPSec authentication header is such an example.

Therefore, in order to solve all kinds of problems existing in IPv4 comprehensively, the next generation Internet Protocol IPv6 designed by IETF has become the only feasible solution at present.

First of all, the 128 bits addressing scheme of IPv6 Protocol can guarantee to provide enough globally unique IP addresses for global IP network nodes in the range of time and space. Moreover, besides increasing address space, IPv6 also enhanced many other essential designs of IPv4.

Hierarchical addressing scheme facilitates Route Aggregation, effectively reduces route table entries and enhances the efficiency and expansibility of routing and data packet processing.

The header design of IPv6 is more efficient compared with IPv4. It has less data fields and takes out header checksum, thus expedites the processing speed of basic IPv6 header. In IPv6 header, fragment field can be shown as an optional extended field, so that data packets fragmentation process won't be done in router forwarding process, and Path MTU Discovery Mechanism collaborates with data packet source which enhances the processing efficiency of router.

Address automatic configuration and plug-and-play is supported. Large amounts of hosts can find network routers easily by address automatic configuration function of IPv6 while obtaining a globally unique IPv6 address automatically as well which makes the devices using IPv6 Internet plug-and-play. Automatic address configuration function also makes the readdressing of existing network easier and more convenient, and it is more convenient for network operators to manage the transformation from one provider to another.

Support IPSec. IPSec is optional in IPv4, but required in IPv6 Protocol. IPv6 provides security extended header, which provides end-to-end security services such as access control, confidentiality and data integrity, consequently making the implement of encryption, validation and Virtual Private Network easier.

Enhance the support for Mobile IP and mobile calculating devices. The Mobile IP Protocol defined in IETF standard makes mobile devices movable without cutting the existing connection, which is a network function getting more and more important. Unlike IPv4, the mobility of IPv6 is from embedded automatic configuration to get transmission address (Care-Of-Address); therefore it doesn't need Foreign Agent. Furthermore, this kind of binding process

enables Correspondent Node communicate with Mobile Node directly, thereby avoids the extra system cost caused by triangle routing choice required in IPv4.

Avoid the use of Network Address Translation. The purpose of the introduction of NAT mechanism is to share and reuse same address space among different network segments. This mechanism mitigates the problem of the shortage of IPv4 address temporally; meanwhile it adds the burden of address translation process for network device and application. Since the address space of IPv6 has increased greatly, address translation becomes unnecessary, thus the problems and system cost caused by NAT deployment are solved naturally.

Support extensively deployed Routing Protocol. IPv6 has kept and extended the supports for existing Internal Gateway Protocols (IGP for short), and Exterior Gateway Protocols (EGP for short). For example, IPv6 Routing Protocol such as RIPng, OSPFv3, IS-ISv6 and MBGP4+, etc.

Multicast addresses increased and the support for multicast has enhanced. By dealing with IPv4 broadcast functions such as Router Discovery and Router Query, IPv6 multicast has completely replaced IPv4 broadcast in the sense of function. Multicast not only saves network bandwidth, but enhances network efficiency as well.

16.2.2 IP Configuration

Layer 3 interface can be configured as IPv4 interface, IPv6 interface.

16.2.2.1 IPv4 Address Configuration

IPv4 address configuration task list:

1. Configure the IPv4 address of three-layer interface

1. Configure the IPv4 address of three-layer interface

Command	Explanation
VLAN Interface Configuration Mode	
ip address *<ip-address>* *<mask>* [secondary] **no ip address [<ip-address> <mask>]**	Configure IP address of VLAN interface; the **no ip address [*<ip-address>* *<mask>*]** command cancels IP address of VLAN interface.

16.2.2.2 IPv6 Address Configuration

The configuration Task List of IPv6 is as follows:

1. IPv6 basic configuration

(1) Configure interface IPv6 address

(2) Configure IPv6 static routing

2. IPv6 Neighbor Discovery Configuration

(1) Configure DAD neighbor solicitation message number
(2) Configure send neighbor solicitation message interval
(3) Enable and disable router advertisement
(4) Configure router lifespan
(5) Configure router advertisement minimum interval
(6) Configure router advertisement maximum interval
(7) Configure prefix advertisement parameters
(8) Configure static IPv6 neighbor entries
(9) Delete all entries in IPv6 neighbor table
(10) Set the hoplimit of sending router advertisement
(11) Set the mtu of sending router advertisement
(12) Set the reachable-time of sending router advertisement
(13) Set the retrans-timer of sending router advertisement
(14) Set the flag representing whether information other than the address information will be obtained via DHCPv6
(15) Set the flag representing whether the address information will be obtained via DHCPv6
3. IPv6 Tunnel configuration
(1) Create/Delete Tunnel
(2) Configure tunnel description
(3) Configure Tunnel Source
(4) Configure Tunnel Destination
(5) Configure Tunnel Next-Hop
(6) Configure Tunnel Mode
(7) Configure Tunnel Routing

1. IPv6 Basic Configuration
(1) Configure interface IPv6 address

Command	Explanation
Interface Configuration Mode	
ipv6 address <ipv6-address/prefix-length> [eui-64] **no ipv6 address <ipv6-address/prefix-length>**	Configure IPv6 address, including aggregatable global unicast addresses, site-local addresses and link-local addresses. The **no ipv6 address <*ipv6-address/prefix-length*>** command cancels IPv6 address.

(2) Set IPv6 Static Routing

Command	Explanation
Global mode	
ipv6 route <ipv6-prefix/prefix-length> {<nexthop-ipv6-address>|<interface-type interface-number> | {<nexthop-ipv6-address> <interface-type interface-number>}} [distance] **no ipv6 route <ipv6-prefix/prefix-length> {<nexthop-ipv6-address>|<interface-type interface-number> |{<nexthop-ipv6-address> <interface-type interface-number>}} [distance]**	Configure IPv6 static routing. The no command cancels IPv6 static routing.

2. IPv6 Neighbor Discovery Configuration

(1) Configure DAD Neighbor solicitation Message number

Command	Explanation
Interface Configuration Mode	
ipv6 nd dad attempts <value> **no ipv6 nd dad attempts**	Set the neighbor query message number sent in sequence when the interface makes duplicate address detection. The no command resumes default value (1).

(2) Configure Send Neighbor solicitation Message Interval

Command	Explanation
Interface Configuration Mode	
ipv6 nd ns-interval <*seconds*> **no ipv6 nd ns-interval**	Set the interval of the interface to send neighbor query message. The NO command resumes default value (1 second).

(3) Enable and disable router advertisement

Command	Explanation
Interface Configuration Mode	
ipv6 nd suppress-ra **no ipv6 nd suppress-ra**	Forbid IPv6 Router Advertisement. The NO command enables IPv6 router advertisement.

(4) Configure Router Lifespan

Command	Explanation
Interface Configuration Mode	
ipv6 nd ra-lifetime <*seconds*> **no ipv6 nd ra-lifetime**	Configure Router advertisement Lifespan. The NO command resumes default value (1800 seconds).

(5) Configure router advertisement Minimum Interval

Command	Description
Interface Configuration Mode	
ipv6 nd min-ra-interval <seconds> **no ipv6 nd min-ra-interval**	Configure the minimum interval for router advertisement. The NO command resumes default value (200 seconds).

(6) Configure router advertisement Maximum Interval

Command	Explanation
Interface Configuration Mode	
ipv6 nd max-ra-interval <*seconds*> **no ipv6 nd max-ra-interval**	Configure the maximum interval for router advertisement. The NO command resumes default value (600 seconds).

(7) Configure prefix advertisement parameters

Command	Explanation
Interface Configuration Mode	
ipv6 nd prefix <ipv6-address/prefix-length> <valid-lifetime> <preferred-lifetime> [off-link] [no-autoconfig] **no ipv6 nd prefix <ipv6-address/prefix-length> <valid-lifetime> <preferred-lifetime> [off-link] [no-autoconfig]**	Configure the address prefix and advertisement parameters of router. The NO command cancels the address prefix of routing advertisement.

(8) Configure static IPv6 neighbor Entries

Command	Explanation
Interface Configuration Mode	
ipv6 neighbor <ipv6-address> <hardware-address> interface	Set static neighbor table entries, including neighbor IPv6 address, MAC address and two-layer port.

<interface-type interface-name>	
no ipv6 neighbor *<ipv6-address>*	Delete neighbor table entries.

(9) Delete all entries in IPv6 neighbor table

Command	Explanation
Admin Mode	
clear ipv6 neighbors	Clear all static neighbor table entries.

(10) Set the hoplimit of sending router advertisement

Command	Explanation
Interface Configuration Mode	
ipv6 nd ra-hoplimit <value>	Set the hoplimit of sending router advertisement.

(11) Set the mtu of sending router advertisement

Command	Explanation
Interface Configuration Mode	
ipv6 nd ra-mtu <value>	Set the mtu of sending router advertisement.

(12) Set the reachable-time of sending router advertisement

Command	Explanation
Interface Configuration Mode	
ipv6 nd reachable-time <seconds>	Set the reachable-time of sending router advertisement.

(13) Set the retrans-timer of sending router advertisement

Command	Explanation
Interface Configuration Mode	
ipv6 nd retrans-timer <seconds>	Set the retrans-timer of sending router advertisement.

(14) Set the flag representing whether information other than the address information will be obtained via DHCPv6.

Command	Explanation
Interface Configuration Mode	
ipv6 nd other-config-flag	Set the flag representing whether information other than the address information will be obtained via DHCPv6.

(15) Set the flag representing whether the address information will be obtained via DHCPv6

Command	Explanation
Interface Configuration Mode	
ipv6 nd managed-config-flag	Set the flag representing whether the address information will be obtained via DHCPv6.

3. IPv6 Tunnel Configuration

(1) Add/Delete tunnel

Command	Explanation
Global mode	
interface tunnel <*tnl-id*> **no interface tunnel <*tnl-id*>**	Create a tunnel. The NO command deletes a tunnel.

(2) Configure tunnel description

Command	Explanation
Tunnel Configuration Mode	
description <*desc*> **no description**	Configure tunnel description. The NO command deletes the tunnel description.

(3) Configure tunnel source

Command	Explanation
Tunnel Configuration Mode	
tunnel source { <ipv4-address> | **<ipv6-address>| <interface-name> }** **no tunnel source**	Configure tunnel source end IPv4/IPv6 address. The NO command deletes the IPv4/IPv6 address of tunnel source end.

(4) Configure Tunnel Destination

Command	Explanation
Tunnel Configuration Mode	
tunnel destination {<ipv4-address> | **<ipv6-address>}** **no tunnel destination**	Configure tunnel destination end IPv4/IPv6 address. The NO command deletes the IPv4/IPv6 address of tunnel destination end.

(5) Configure Tunnel Next-Hop

Command	Explanation
Tunnel Configuration Mode	

tunnel nexthop <ipv4-address> **no tunnel nexthop**	Configure tunnel next-hop IPv4 address. The NO command deletes the IPv4 address of tunnel next-hop end.

(6) Configure Tunnel Mode

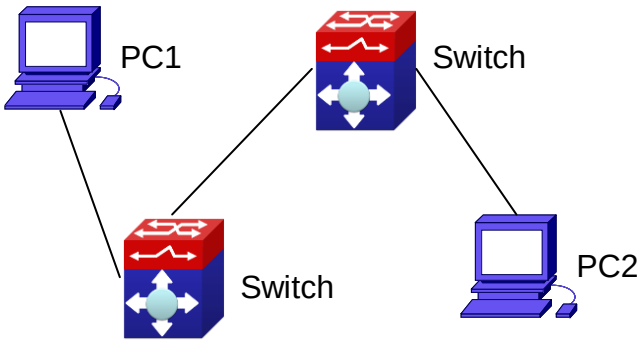
Command	Explanation
Tunnel Configuration Mode	
tunnel mode [[gre] | ipv6ip [6to4 | isatap]] **no tunnel mode**	Configure tunnel mode. The NO command clears tunnel mode.

(7) Configure Tunnel Routing

Command	Explanation
Global mode	
ipv6 route <ipv6-address/prefix-length> {<interface-type interface-number> | tunnel <tnl-id>} **no ipv6 route <ipv6-address/prefix-length> {<interface-type interface-number> | tunnel <tnl-id>}**	Configure tunnel routing. The NO command clears tunnel routing.

16.2.3 IP Configuration Examples

16.2.3.1 Configuration Examples of IPv4



IPv4 configuration example

The user's configuration requirements are: Configure IP address of different network segments on Switch1 and Switch2, configure static routing and validate accessibility using ping function.

Configuration Description:

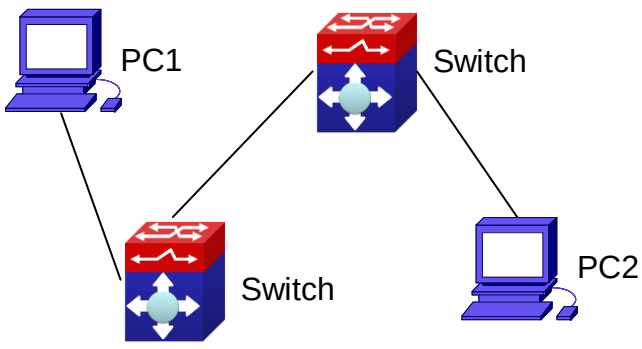
Configure two VLANs on Switch1, namely, VLAN1 and VLAN2.

Configure IPv4 address 192.168.1.1 255.255.255.0 in VLAN1 of Switch1, and configure IPv4 address 192.168.2.1 255.255.255.0 in VLAN2.
Configure two VLANs on Switch2, respectively VLAN2 and VLAN3.
Configure IPv4 address 192.168.2.2 255.255.255.0 in VLAN2 of Switch2, and configure IPv4 address 192.168.3.1 255.255.255.0 in VLAN3.
The IPv4 address of PC1 is 192.168.1.100 255.255.255.0, and the IPv4 address of PC2 is 192.168.3.100 255.255.255.0.
Configure static routing 192.168.3.0/24 on Switch1, and configure static routing 192.168.1.0/24 on Switch2.
Ping each other among PCs.
Note: First make sure PC1 and Switch1 can access each other by ping, and PC2 and Switch2 can access each other by ping.
The configuration procedure is as follows:

```
Switch1(config)#interface vlan 1
Switch1(Config-if-Vlan1)#ip address 192.168.1.1 255.255.255.0
Switch1(config)#interface vlan 2
Switch1(Config-if-Vlan2)#ip address 192.168.2.1 255.255.255.0
Switch1(Config-if-Vlan2)#exit
Switch1(config)#ip route 192.168.3.0 255.255.255.0 192.168.2.2

Switch2(config)#interface vlan 2
Switch2(Config-if-Vlan2)#ip address 192.168.2.2 255.255.255.0
Switch2(config)#interface vlan 3
Switch2(Config-if-Vlan3)#ip address 192.168.3.1 255.255.255.0
Switch2(Config-if-Vlan3)#exit
Switch2(config)#ip route 192.168.1.0 255.255.255.0 192.168.2.1
```

16.2.3.2 Configuration Examples of IPv6

Example 1:



IPv6 configuration example

The user's configuration requirements are: Configure IPv6 address of different network segments on Switch1 and Switch2, configure static routing and validate reachability using ping6 function.

Configuration Description:

Configure two VLANs on Switch1, namely, VLAN1 and VLAN2.

Configure IPv6 address 2001::1/64 in VLAN1 of Switch1, and configure IPv6 address 2002::1/64 in VLAN2.

Configure 2 VLANs on Switch2, namely, VLAN2 and VLAN3.

Configure IPv6 address 2002::2/64 in VLAN2 of Switch2, and configure IPv6 address 2003::1/64 in VLAN3.

The IPv6 address of PC1 is 2001::11/64, and the IPv6 address of PC2 is 2003::33/64.

Configure static routing 2003:33/64 on Switch1, and configure static routing 2001::11/64 on Switch2.

ping6 each other among PCs.

Note: First make sure PC1 and Switch1 can access each other by ping, and PC2 and Switch2 can access each other by ping.

The configuration procedure is as follows:

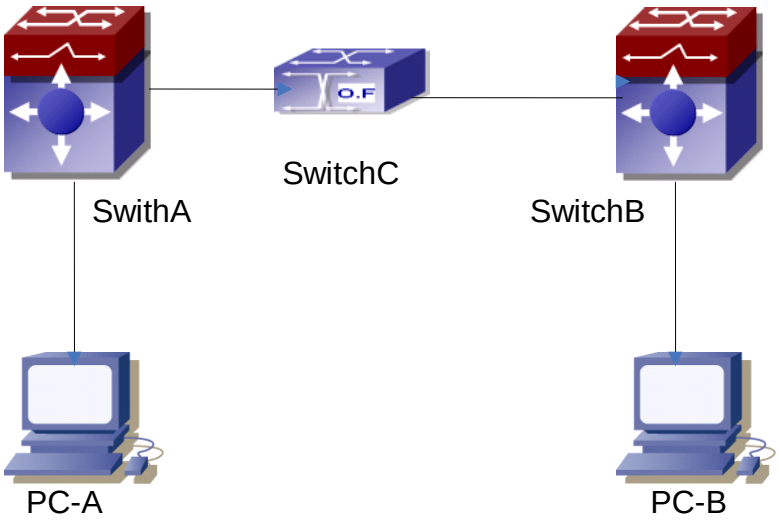
```
Switch1(Config)#interface vlan 1
Switch1(Config-if-Vlan1)#ipv6 address 2001::1/64
Switch1(Config)#interface vlan 2
Switch1(Config-if-Vlan2)#ipv6 address 2002::1/64
Switch1(Config-if-Vlan2)#exit
Switch1(Config)#ipv6 route 2003::33/64 2002::2

Switch2(Config)#interface vlan 2
Switch2(Config-if-Vlan2)#ipv6 address 2002::2/64
Switch2(Config)#interface vlan 3
Switch2(Config-if-Vlan3)#ipv6 address 2003::1/64
Switch2(Config-if-Vlan3)#exit
Switch2(Config)#ipv6 route 2001::33/64 2002::1

Switch1#ping6 2003::33
```

Configuration result:

```
Switch1#show run
interface Vlan1
 ipv6 address 2001::1/64
!
interface Vlan2
 ipv6 address 2002::2/64
!
```

```
interface Loopback
 mtu 3924
!
ipv6 route 2003::/64 2002::2
!
no login
!
end

Switch2#show run
interface Vlan2
 ipv6 address 2002::2/64
!
interface Vlan3
 ipv6 address 2003::1/64
!
interface Loopback
 mtu 3924
!
ipv6 route 2001::/64 2002::1
!
no login
!
End
```

Example 2:



This case is IPv6 tunnel with the following user configuration requirements: SwitchA and SwitchB are tunnel nodes, dual-stack is supported. SwitchC only runs IPv4, PC-A and PC-B communicate.

Configuration Description:

Configure two vlans on SwitchA, namely, VLAN1 and VLAN2. VLAN1 is IPv6 domain, VLAN2 connects to IPv4 domain.

Configure IPv6 address 2002:caca:ca01:2::1/64 in VLAN1 of SwitchA and turn on RA function, configure IPv4 address 202.202.202.1 in VLAN2.

Configure two VLANs on SwitchB, namely, VLAN3 and VLAN4, VLAN4 is IPv6 domain, and VLAN3 connects to IPv4 domain.

Configure IPv6 address 2002:cbcb:cb01:2::1/64 in VLAN4 of SwitchB and turn on RA function, configure IPv4 address 203.203.203.1 on VLAN3.

Configure tunnel on SwitchA, the source IPv4 address of the tunnel is 202.202.202.1, the tunnel routing is ::/0

Configure tunnel on SwitchB, the source IPv4 address of the tunnel is 203.203.203.1, and the tunnel routing is ::/0

Configure two VLANs on SwitchC, namely, VLAN2 and VLAN3. Configure IPv4 address 202.202.202.202 on VLAN2 and configure IPv4 address 203.203.203.203 on VLAN3.

PC-A and PC-B get the prefix of 2002 via SwitchA and SwitchB to configure IPv6 address automatically.

On PC-A, ping IPv6 address of PC-B

The configuration procedure is as follows:

```
SwitchA(Config-if-Vlan1)#ipv6 address 2002:caca:ca01:2::1/64
SwitchA(Config-if-Vlan1)#no ipv6 nd suppress-ra
SwitchA(Config-if-Vlan1)#interface vlan 2
SwitchA(Config-if-Vlan2)#ipv4 address 202.202.202.1 255.255.255.0
SwitchA(Config-if-Vlan1)#exit
SwitchA(config)# interface tunnel 1
SwitchA(Config-if-Tunnel1)#tunnel source 202.202.202.1
SwitchA(Config-if-Tunnel1)#tunnel destination 203.203.203.1
SwitchA(Config-if-Tunnel1)#tunnel mode ipv6ip
SwitchA(config)#ipv6 route ::/0 tunnel1

SwitchB(Config-if-Vlan4)#ipv6 address 2002:cbcb:cb01::2/64
SwitchB(Config-if-Vlan4)#no ipv6 nd suppress-ra
SwitchB (Config-if-Vlan3)#interface vlan 3
SwitchB (Config-if-Vlan2)#ipv4 address 203.203.203.1 255.255.255.0
SwitchB (Config-if-Vlan1)#exit
```

SwitchB(config)#interface tunnel 1
SwitchB(Config-if-Tunnel1)#tunnel source 203.203.203.1
SwitchB(Config-if-Tunnel1)#tunnel destination 202.202.202.1
SwitchB(Config-if-Tunnel1)#tunnel mode ipv6ip
SwitchB(config)#ipv6 route ::/0 tunnel1

16.2.4 IPv6 Troubleshooting

The router lifespan configured should not be smaller than the Send Router advertisement Interval. If the connected PC has not obtained IPv6 address, you should check RA announcement switch (the default is turned off).

16.3 IP Forwarding

16.3.1 Introduction to IP Forwarding

Gateway devices can forward IP packets from one subnet to another; such forwarding uses routes to find a path. IP forwarding of switch is done with the participation of hardware, and can achieve wire speed forwarding. In addition, flexible management is provided to adjust and monitor forwarding. Switch supports aggregation algorithm enabling/disabling optimization to adjust generation of network route entry in the switch chip and view statistics for IP forwarding and hardware forwarding chip status.

16.3.2 IP Route Aggregation Configuration Task

IP route aggregation configuration task:

1. Set whether IP route aggregation algorithm with/without optimization should be used

Command	Explanation
Global Mode	
ip fib optimize **no ip fib optimize**	Enables the switch to use optimized IP route aggregation algorithm; the "**no ip fib optimize**" disables the optimized IP route aggregation algorithm.

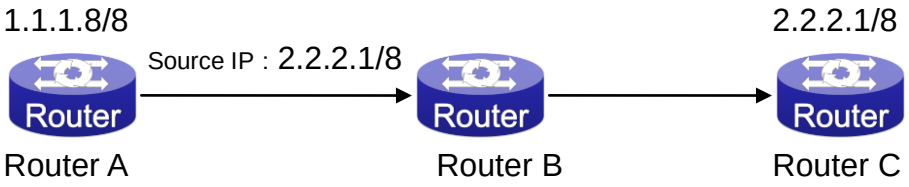
16.4 URPF

16.4.1 Introduction to URPF

URPF (Unicast Reverse Path Forwarding) introduces the RPF technology applied in multicast to unicast, so to protect the network from the attacks which is based on source address cheat. When switch receives the packet, it will search the route in the route table using the source

address as the destination address which is acquired from the packet. If the found router exit interface does not match the entrance interface acquired from this packet, the switch will consider this packet a fake packet and discard it.

In Source Address Spoofing attacks, attackers will construct a series of messages with fake source addresses. For applications based on IP address verification, such attacks may allow unauthorized users to access the system as some authorized ones, or even the administrator. Even if the response messages can't reach the attackers, they will also damage the targets.



URPF application situation

In the above figure, Router A sends requests to the server Router B by faking messages whose source address are 2.2.2.1/8 .In response, Router B will send the messages to the real "2.2.2.1/8". Such illegal messages attack both Router B and Router C. The application of URPF technology in the situation described above can avoid the attacks based on the Source Address Spoofing.

16.4.2 URPF Configuration Task Sequence

Enable URPF

Display and debug URPF relevant information

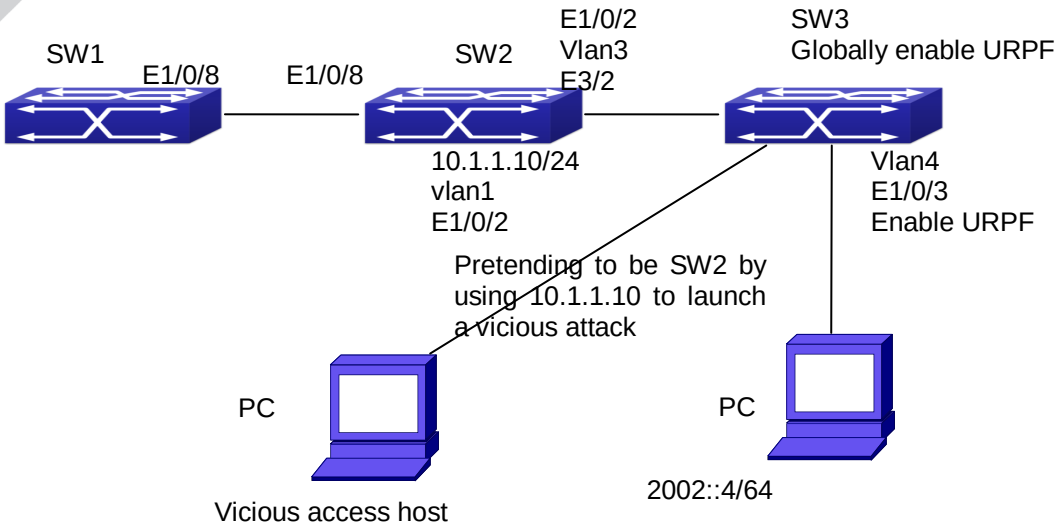
1. Globally enable URPF

Command	Explanation
Global mode	
urpf enable **no urpf enable**	Globally enable and disable URPF.

2. Display and debug URPF relevant information

Command	Explanation
Admin and Config Mode	
show urpf	Display which interfaces have been enabled with URPF function.

16.4.3 URPF Typical Example



In the network, topology shown in the graph above, IP URPF function is enabled on SW3. When there is someone in the network pretending to be someone else by using his IP address to launch a vicious attack, the switch will drop all the attacking messages directly through the hardware function.

Enable the URPF function in SW3.

SW3 configuration task sequence:

Switch3#config

Switch3(config)#urpf enable

16.4.4 URPF Troubleshooting

If all configurations are normal but URPF still can't operate as expected, please enable the URPF debug function and use "show urpf" command to observe whether URPF is enabled, and send the result to the technology service center.

16.5 ARP

16.5.1 Introduction to ARP

ARP (Address Resolution Protocol) is mainly used to resolve IP address to Ethernet MAC address. Switch supports both dynamic ARP and static ARP configuration.Furthermore, switch supports the configuration of proxy ARP for some applications. For instance, when an ARP request is received on the port, requesting an IP address in the same IP segment of the port but not the same physical network, if the port has enabled proxy ARP, the port would reply to the ARP with its own MAC address and forward the actual packets received. Enabling proxy ARP allows machines physically separated but of the same IP segment ignores the physical

separation and communicate via proxy ARP interface as if in the same physical network.

16.5.2 ARP Configuration Task List

ARP Configuration Task List:

1. Configure static ARP
2. Configure proxy ARP
3. Clear dynamic ARP
4. Clear the statistic information of ARP messages

1. Configure static ARP

Command	Explanation
VLAN Interface Mode	
arp <*ip_address*> <*mac_address*> {interface [ethernet] <portName>} **no arp <ip_address>**	Configures a static ARP entry; the no command deletes a ARP entry of the specified IP address.

2. Configure proxy ARP

Command	Explanation
VLAN Interface Mode	
ip proxy-arp **no ip proxy-arp**	Enables the proxy ARP function for Ethernet ports: the no command disables the proxy ARP.

3. Clear dynamic ARP

Command	Explanation
Admin mode	
clear arp-cache	Clear the dynamic ARP learnt by the switch.

4. Clear the statistic information of ARP message

Command	Explanation
Admin mode	
clear arp traffic	Clear the statistic information of ARP messages of the switch.

16.5.3 ARP Troubleshooting

If ping from the switch to directly connected network devices fails, the following can be used to check the possible cause and create a solution.

Check whether the corresponding ARP has been learned by the switch.
If ARP has not been learned, then enabled ARP debugging information and view the sending/receiving condition of ARP packets.
Defective cable is a common cause of ARP problems and may disable ARP learning.

16.6 Hardware Tunnel Capacity Configuration

16.6.1 Introduction to Hardware Tunnel Capacity

Hardware Tunnel Capacity is the maximum number of tunnel and MPLS forwarded by hardware. Capacity can be adjusted by this command, increasing capacity will reduce hardware routing number supported by switch, vice versa.

16.6.2 Hardware Tunnel Capacity Configuration

Hardware Tunnel Capacity Configuration Task List:
1. Configure hardware tunnel capacity

Command	Explanation
Global mode	
hardware tunnel-capacity < *size*> **no hardware tunnel-capacity**	Configure capacity of hardware tunnel and MPLS, the no command restores the default capacity.

Note: after adjust hardware tunnel capacity, it needs to reset switch to enable the valid configuration.

16.6.3 Hardware Tunnel Capacity Troubleshooting

After adjust hardware tunnel capacity, it must save the configuration and reset switch, the configuration can takes effect.

Chapter 17 ARP Scanning Prevention Function Configuration

17.1 Introduction to ARP Scanning Prevention Function

ARP scanning is a common method of network attack. In order to detect all the active hosts in a network segment, the attack source will broadcast lots of ARP messages in the segment, which will take up a large part of the bandwidth of the network. It might even do large-traffic-attack in the network via fake ARP messages to collapse of the network by exhausting the bandwidth. Usually ARP scanning is just a preface of other more dangerous attack methods, such as automatic virus infection or the ensuing port scanning, vulnerability scanning aiming at stealing information, distorted message attack, and DOS attack, etc.

Since ARP scanning threatens the security and stability of the network with great danger, so it is very significant to prevent it. Switch provides a complete resolution to prevent ARP scanning: if there is any host or port with ARP scanning features is found in the segment, the switch will cut off the attack source to ensure the security of the network.

There are two methods to prevent ARP scanning: port-based and IP-based. The port-based ARP scanning will count the number to ARP messages received from a port in a certain time range, if the number is larger than a preset threshold, this port will be “down”. The IP-based ARP scanning will count the number to ARP messages received from an IP in the segment in a certain time range, if the number is larger than a preset threshold, any traffic from this IP will be blocked, while the port related with this IP will not be “down”. These two methods can be enabled simultaneously. After a port or an IP is disabled, users can recover its state via automatic recovery function.

To improve the effect of the switch, users can configure trusted ports and IP, the ARP messages from which will not be checked by the switch. Thus the load of the switch can be effectively decreased.

17.2 ARP Scanning Prevention Configuration Task Sequence

Enable the ARP Scanning Prevention function.

Configure the threshold of the port-based and IP-based ARP Scanning Prevention

Configure trusted ports

Configure trusted IP

Configure automatic recovery time

Display relative information of debug information and ARP scanning

1. Enable the ARP Scanning Prevention function.

Command	Explanation
Global configuration mode	
anti-arpscan enable **no anti-arpscan enable**	Enable or disable the ARP Scanning Prevention function globally.

2. Configure the threshold of the port-based and IP-based ARP Scanning Prevention

Command	Explanation
Global configuration mode	
anti-arpscan port-based threshold <threshold-value> **no anti-arpscan port-based threshold**	Set the threshold of the port-based ARP Scanning Prevention.
anti-arpscan ip-based threshold *<threshold-value>* **no anti-arpscan ip-based threshold**	Set the threshold of the IP-based ARP Scanning Prevention.

3. Configure trusted ports

Command	Explanation
Port configuration mode	
anti-arpscan trust <port | supertrust-port> **no anti-arpscan trust <port | supertrust-port>**	Set the trust attributes of the ports.

4. Configure trusted IP

Command	Explanation
Global configuration mode	
anti-arpscan trust ip *<ip-address>* [*<netmask>*] **no anti-arpscan trust ip *<ip-address>* [*<netmask>*]**	Set the trust attributes of IP.

5. Configure automatic recovery time

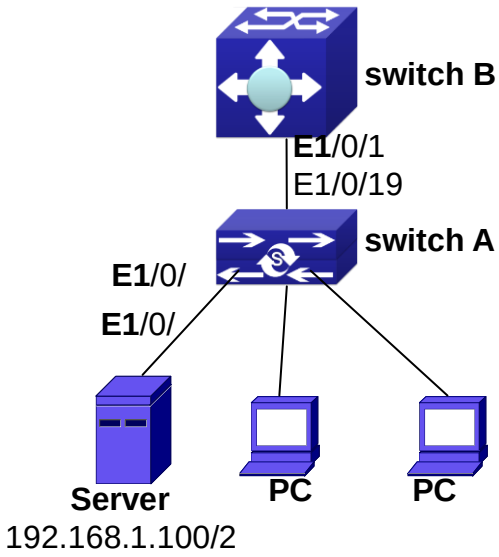
Command	Explanation
Global configuration mode	
anti-arpscan recovery enable **no anti-arpscan recovery enable**	Enable or disable the automatic recovery function.
anti-arpscan recovery time *<seconds>*	Set automatic recovery time.

no anti-arpscan recovery time	

Display relative information of debug information and ARP scanning

Command	Explanation
Global configuration mode	
anti-arpscan log enable **no anti-arpscan log enable**	Enable or disable the log function of ARP scanning prevention.
anti-arpscan trap enable **no anti-arpscan trap enable**	Enable or disable the SNMP Trap function of ARP scanning prevention.
show anti-arpscan [trust <ip | port | supertrust-port> | prohibited <ip | port>]	Display the state of operation and configuration of ARP scanning prevention.
Admin Mode	
debug anti-arpscan <*port* | *ip*> **no debug anti-arpscan <*port* | *ip*>**	Enable or disable the debug switch of ARP scanning prevention.

17.3 ARP Scanning Prevention Typical Examples



ARP scanning prevention typical configuration example

In the network topology above, port E1/0/1 of switch B is connected to port E1/0/19 of switch A, the port E1/0/2 of switch A is connected to file server (IP address is 192.168.1.100/24), and all the other ports of switch A are connected to common PC. The following configuration can prevent ARP scanning effectively without affecting the normal operation of the system.

switch A configuration task sequence:

SwitchA(config)#anti-arpscan enable

SwitchA(config)#anti-arpscan recovery time 3600

SwitchA(config)#anti-arpscan trust ip 192.168.1.100 255.255.255.0

SwitchA(config)#interface ethernet1/0/2
SwitchA (Config-If-Ethernet1/0/2)#anti-arpscan trust port
SwitchA (Config-If-Ethernet1/0/2)#exit
SwitchA(config)#interface ethernet1/0/19
SwitchA (Config-If-Ethernet1/0/19)#anti-arpscan trust supertrust-port
Switch A(Config-If-Ethernet1/0/19)#exit

switchB configuration task sequence:
Switch B(config)# anti-arpscan enable
SwitchB(config)#interface ethernet1/0/1
SwitchB (Config-If-Ethernet 1/0/1)#anti-arpscan trust port
SwitchB (Config-If-Ethernet 1/0/1)exit

17.4 ARP Scanning Prevention Troubleshooting Help

ARP scanning prevention is disabled by default. After enabling ARP scanning prevention, users can enable the debug switch, “**debug anti-arpscan**”, to view debug information.

Chapter 18 Prevent ARP, ND Spoofing Configuration

18.1 Overview

18.1.1 ARP (Address Resolution Protocol)

Generally speaking, ARP (RFC-826) protocol is mainly responsible of mapping IP address to relevant 48-bit physical address, that is MAC address, for instance, IP address is 192.168.0.1, network card Mac address is 00-1F-CE-FD-1D-2B. What the whole mapping process is that a host computer send broadcast data packet involving IP address information of destination host computer, ARP request, and then the destination host computer send a data packet involving its IP address and Mac address to the host, so two host computers can exchange data by MAC address.

18.1.2 ARP Spoofing

In terms of ARP Protocol design, to reduce redundant ARP data communication on networks, even though a host computer receives an ARP reply which is not requested by itself, it will also insert an entry to its ARP cache table, so it creates a possibility of “ARP spoofing”. If the hacker wants to snoop the communication between two host computers in the same network (even if are connected by the switches), it sends an ARP reply packet to two hosts separately, and make them misunderstand MAC address of the other side as the hacker host MAC address. In this way, the direct communication is actually communicated indirectly among the hacker host computer. The hackers not only obtain communication information they need, but also only need to modify some information in data packet and forward successfully. In this sniff way, the hacker host computer doesn’t need to configure intermix mode of network card, that is because the data packet between two communication sides are sent to hacker host computer on physical layer, which works as a relay.

18.1.3 How to prevent void ARP/ND Spoofing

There are many sniff, monitor and attack behaviors based on ARP protocol in networks, and most of attack behaviors are based on ARP spoofing, so it is very important to prevent ARP spoofing. ARP spoofing accesses normal network environment by counterfeiting legal IP address firstly, and sends a great deal of counterfeited ARP application packets to switches, after switches learn these packets, they will cover previously corrected IP, mapping of MAC address, and then some corrected IP, MAC address mapping are modified to correspondence relationship configured by attack packets so that the switch makes mistake on transfer packets, and takes an effect on the whole network. Or the switches are made used of by vicious attackers, and they intercept and capture packets transferred by switches or attack other

switches, host computers or network equipment.

What the essential method on preventing attack and spoofing switches based on ARP in networks is to disable switch automatic update function; the cheater can't modify corrected MAC address in order to avoid wrong packets transfer and can't obtain other information. At one time, it doesn't interrupt the automatic learning function of ARP. Thus it prevents ARP spoofing and attack to a great extent.

ND is neighbor discovering protocol in IPv6 protocol, and it's similar to ARP on operation principle, therefore we do in the same way as preventing ARP spoofing to prevent ND spoofing and attack.

18.2 Prevent ARP, ND Spoofing configuration

The steps of preventing ARP, ND spoofing configuration as below:

Disable ARP, ND automatic update function

Disable ARP, ND automatic learning function

Changing dynamic ARP, ND to static ARP, ND

1. Disable ARP, ND automatic update function

Command	Explanation
Global Mode and Port Mode	
ip arp-security updateprotect **no ip arp-security updateprotect** **ipv6 nd-security updateprotect** **no ipv6 nd-security updateprotect**	Disable and enable ARP, ND automatic update function.

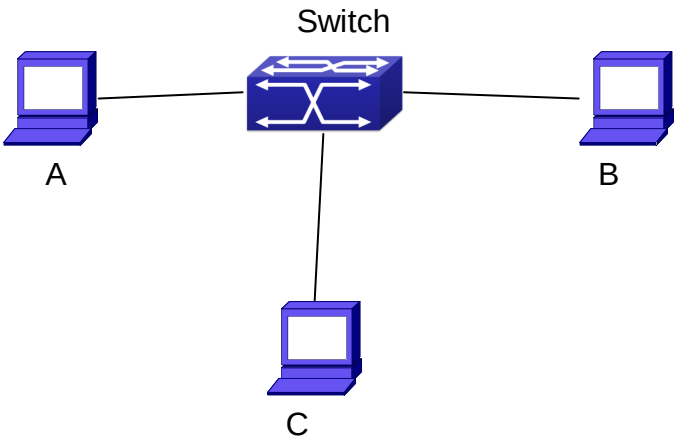
2. Disable ARP, ND automatic learning function

Command	Explanation
Global mode and Interface Mode	
ip arp-security learnprotect **no ip arp-security learnprotect** **ipv6 nd-security learnprotect** **no ipv6 nd-security learnprotect**	Disable and enable ARP, ND automatic learning function.

3. Function on changing dynamic ARP, ND to static ARP, ND

Command	Explanation
Global Mode and Port Mode	
ip arp-security convert **ipv6 nd-security convert**	Change dynamic ARP, ND to static ARP, ND.

18.3 Prevent ARP, ND Spoofing Example



Equipment Explanation

Equipment	Configuration	Quality
switch	IP:192.168.2.4; IP:192.168.1.4; mac: 00-00-00-00-00-04	1
A	IP:192.168.2.1; mac: 00-00-00-00-00-01	1
B	IP:192.168.1.2; mac: 00-00-00-00-00-02	1
C	IP:192.168.2.3; mac: 00-00-00-00-00-03	some

There is a normal communication between B and C on above diagram. A wants switch to forward packets sent by B to itself, so need switch sends the packets transfer from B to A. firstly A sends ARP reply packet to switch, format is: 192.168.2.3, 00-00-00-00-00-01, mapping its MAC address to C's IP, so the switch changes IP address when it updates ARP list., then data packet of 192.168.2.3 is transferred to 00-00-00-00-00-01 address (A MAC address).

In further, a transfers its received packets to C by modifying source address and destination address, the mutual communicated data between B and C are received by A unconsciously. Because the ARP list is update timely, another task for A is to continuously send ARP reply packet, and refreshes switch ARP list.

So it is very important to protect ARP list, configure to forbid ARP learning command in stable environment, and then change all dynamic ARP to static ARP, the learned ARP will not be refreshed, and protect for users.

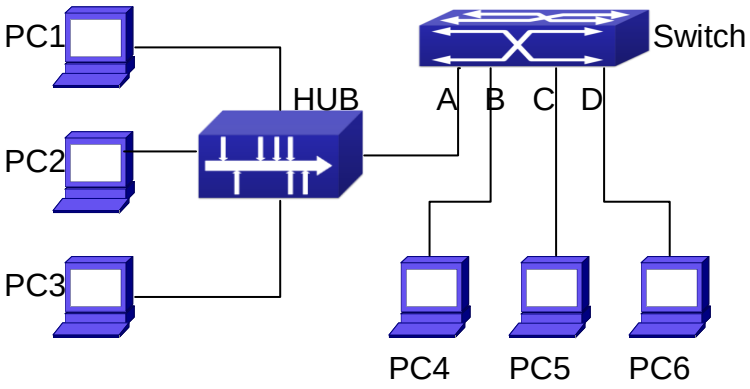
Switch#config
Switch(config)#interface vlan 1
Switch(Config-If-Vlan1)#arp 192.168.2.1 00-00-00-00-00-01 interface eth 1/0/2
Switch(Config-If-Vlan1)#interface vlan 2
Switch(Config-If-Vlan2)#arp 192.168.1.2 00-00-00-00-00-02 interface eth 1/0/2
Switch(Config-If-Vlan2#interface vlan 3
Switch(Config-If-Vlan3)#arp 192.168.2.3 00-00-00-00-00-03 interface eth 1/0/2
Switch(Config-If-Vlan3)#exit
Switch(Config)#ip arp-security learnprotect

Switch(Config)#
Switch(config)#ip arp-security convert
If the environment changing, it enable to forbid ARP refresh, once it learns ARP property, it wont be refreshed by new ARP reply packet, and protect use data from sniffing.
Switch#config
Switch(config)#ip arp-security updateprotect

Chapter 19 ARP GUARD Configuration

19.1 Introduction to ARP GUARD

There is serious security vulnerability in the design of ARP protocol, which is any network device, can send ARP messages to advertise the mapping relationship between IP address and MAC address. This provides a chance for ARP cheating. Attackers can send ARP REQUEST messages or ARP REPLY messages to advertise a wrong mapping relationship between IP address and MAC address, causing problems in network communication. The danger of ARP cheating has two forms: 1. PC4 sends an ARP message to advertise that the IP address of PC2 is mapped to the MAC address of PC4, which will cause all the IP messages to PC2 will be sent to PC4, thus PC4 will be able to monitor and capture the messages to PC2; 2. PC4 sends ARP messages to advertise that the IP address of PC2 is mapped to an illegal MAC address, which will prevent PC2 from receiving the messages to it. Particularly, if the attacker pretends to be the gateway and do ARP cheating, the whole network will be collapsed.



ARP GUARD schematic diagram

We utilize the filtering entries of the switch to protect the ARP entries of important network devices from being imitated by other devices. The basic theory of doing this is that utilizing the filtering entries of the switch to check all the ARP messages entering through the port, if the source address of the ARP message is protected, the messages will be directly dropped and will not be forwarded.

ARP GUARD function is usually used to protect the gateway from being attacked. If all the accessed PCs in the network should be protected from ARP cheating, then a large number of ARP GUARD address should be configured on the port, which will take up a big part of FFP entries in the chip, and as a result, might affect other applications. So this will be improper. It is recommended that adopting FREE RESOURCE related accessing scheme. Please refer to relative documents for details.

19.2 ARP GUARD Configuration Task List

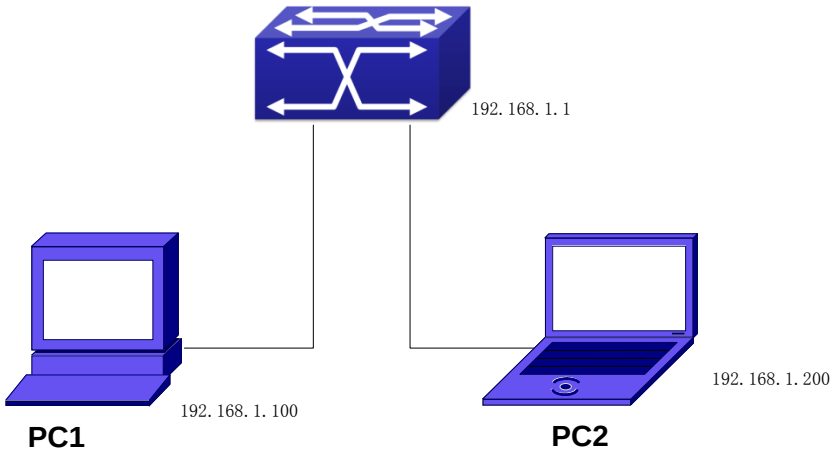
1. Configure the protected IP address

Command	Explanation
Port configuration mode	
arp-guard ip <addr> **no arp-guard ip <addr>**	Configure/delete ARP GUARD address

Chapter 20 ARP Local Proxy Configuration

20.1 Introduction to ARP Local Proxy function

In a real application environment, the switches in the aggregation layer are required to implement local ARP proxy function to avoid ARP cheating. This function will restrict the forwarding of ARP messages in the same vlan and thus direct the L3 forwarding of the data flow through the switch.



As shown in the figure above, PC1 wants to send an IP message to PC2, the overall procedure goes as follows (some non-arp details are ignored)

1. Since PC1 does not have the ARP of PC2, it sends and broadcasts ARP request.
2. Receiving the ARP message, the switch hardware will send the ARP request to CPU instead of forwarding this message via hardware, according to new ARP handling rules.
3. With local ARP proxy enabled, the switch will send ARP reply message to PC1 (to fill up its mac address)
4. After receiving the ARP reply, PC1 will create ARP, send an IP message, and set the destination MAC of the Ethernet head as the MAC of the switch.
5. After receiving the ip message, the switch will search the router table (to create router cache) and distribute hardware entries.
6. If the switch has the ARP of PC2, it will directly encapsulate the Ethernet head and send the message (the destination MAC is that of PC2)
7. If the switch does not have the ARP of PC2, it will request it and then send the ip message.

This function should cooperate with other security functions. When users configure local ARP proxy on an aggregation switch while configuring interface isolation function on the layer-2 switch connected to it, all ip flow will be forwarded on layer 3 via the aggregation switch. And due to the interface isolation, ARP messages will not be forwarded within the vlan, which means other PCs will not receive it.

20.2 ARP Local Proxy Function Configuration Task List

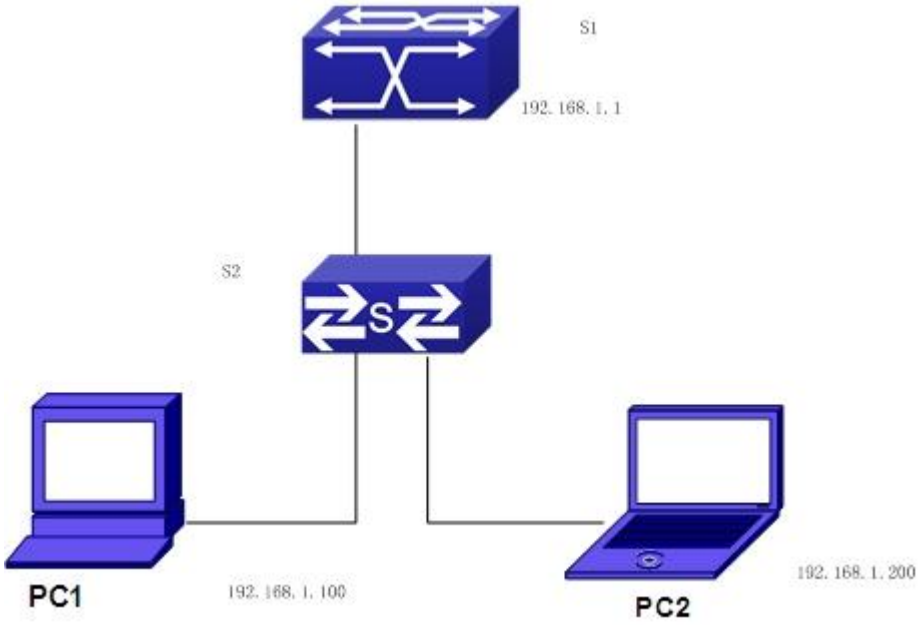
1. Enable/disable ARP local proxy function

Command	Explanation
Interface vlan mode	
ip local proxy-arp **no ip local proxy-arp**	Enable or disable ARP local proxy function.

20.3 Typical Examples of ARP Local Proxy Function

As shown in the following figure, S1 is a medium/high-level layer-3 switch supporting ARP local proxy, S2 is layer-2 access switches supporting interface isolation.

Considering security, interface isolation function is enabled on S2. Thus all downlink ports of S2 is isolated from each other, making all ARP messages able to be forwarded through S1. If ARP local proxy is enabled on S1, then all interfaces on S1 isolate ARP while S1 serves as an ARP proxy. As a result, IP flow will be forwarded at layer 3 through S1 instead of S2.



We can configure as follows:

Switch(config)#interface vlan 1

Switch(Config-if-Vlan1)#ip address 192.168.1.1 255.255.255.0

Switch(Config-if-Vlan1)#ip local proxy-arp

Switch(Config-if-Vlan1)#exit

20.4 ARP Local Proxy Function Troubleshooting

ARP local proxy function is disabled by default. Users can view the current configuration with display command. With correct configuration, by enabling debug of ARP, users can check whether the ARP proxy is normal and send proxy ARP messages.

In the process of operation, the system will show corresponding prompts if any operational error occurs.

Chapter 21 Gratuitous ARP Configuration

21.1 Introduction to Gratuitous ARP

Gratuitous ARP is a kind of ARP request that is sent by the host with its IP address as the destination of the ARP request.

The basic working mode for QTECH switches is as below: The Layer 3 interfaces of the switch can be configured to advertise gratuitous ARP packets period or the switch can be configured to enable to send gratuitous ARP packets in all the interfaces globally.

The purpose of gratuitous ARP is as below:

To reduce the frequency that the host sends ARP request to the switch. The hosts in the network will periodically send ARP requests to the gateway to update the MAC address of the gateway. If the switch advertises gratuitous ARP requests, the host will not have to send these requests. This will reduce the frequency the hosts' sending ARP requests for the gateway's MAC address.

Gratuitous ARP is a method to prevent ARP cheating. The switch's advertising gratuitous ARP request will force the hosts to update its ARP table cache. Thus, forged ARP of gateway cannot function.

21.2 Gratuitous ARP Configuration Task List

Enable gratuitous ARP and configure the interval to send gratuitous ARP request

Display configurations about gratuitous ARP

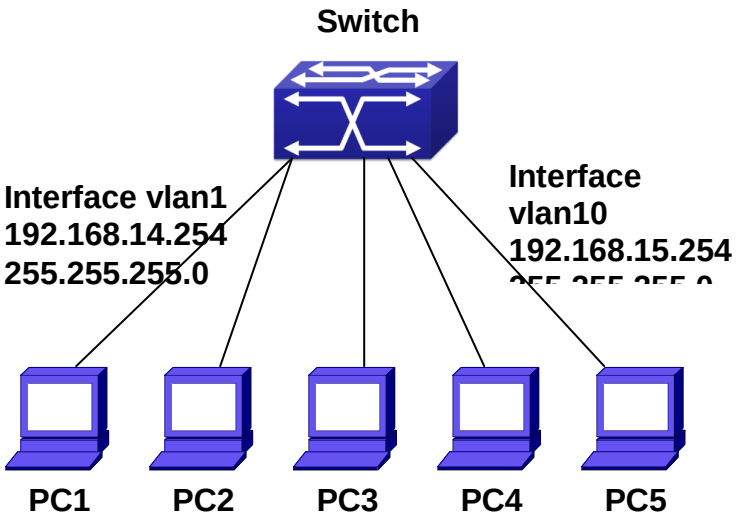
1. Enable gratuitous ARP and configure the interval to send gratuitous ARP request.

Command	Explanation
Global Configuration Mode and Interface Configuration Mode.	
ip gratuitous-arp <*5-1200*> **no ip gratuitous-arp**	To enable gratuitous ARP and configure the interval to send gratuitous ARP request. The no command cancels the gratuitous ARP.

2. Display configurations about gratuitous ARP

Command	Explanation
Admin Mode and Configuration Mode	
show ip gratuitous-arp [interface vlan <*1-4094*>]	To display configurations about gratuitous ARP.

21.3 Gratuitous ARP Configuration Example



Gratuitous ARP Configuration Example

For the network topology shown in the figure above, interface VLAN10 whose IP address is 192.168.15.254 and network address mask is 255.255.255.0 in the switch system. Three PCs – PC3, PC4, PC5 are connected to the interface. The IP address of interface VLAN 1 is 192.168.14.254, its network address mask is 255.255.255.0. Two PCs – PC1 and PC2 are connected to this interface. Gratuitous ARP can be enabled through the following configuration:

Configure two interfaces to use gratuitous ARP at one time.
Switch(config)#ip gratuitous-arp 300
Switch(config)#exit

Configure gratuitous ARP specifically for only one interface at one time.
Switch(config)#interface vlan 10
Switch(Config-if-Vlan10)#ip gratuitous-arp 300
Switch(Config-if-Vlan10)#exit
Switch(config) #exit

21.4 Gratuitous ARP Troubleshooting

Gratuitous ARP is disabled by default. And when gratuitous ARP is enabled, the debugging information about ARP packets can be retrieved through the command debug ARP send.
If gratuitous ARP is enabled in global configuration mode, it can be disabled only in global configuration mode. If gratuitous ARP is configured in interface configuration mode, the configuration can only be disabled in interface configuration mode.

Chapter 22 Keepalive Gateway Configuration

22.1 Introduction to Keepalive Gateway

Ethernet port is used to process backup or load balance, for the reason that it is a broadcast channel, it may not detect the change of physical signal and fails to get to down when the gateway is down. Keepalive Gateway is introduced to detect the connectivity to the higher-up gateway, in the case that a Ethernet port connect with a higher-up gateway to form a point-to-point network topology.

For example: router connects optical terminal device and the line is up all the time, While the line between moden and remote gateway is down, it is necessary to use a effective method to detect whether the remote gateway is reachable. At present, detect gateway connectivity by sending ARP request to gateway on time, if ARP resolution is failing, shutdown the interface, if ARP resolution is successful, keep the interface up.

Only layer 3 switch supports keepalive gateway function.

22.2 Keepalive Gateway Configuration Task List

Enable or disable keepalive gateway, configure the interval period that ARP request packet is sent and the retry-count after detection is failing

Show keepalive gateway and IPv4 running status of the interface

1. Enable or disable keepalive gateway, configure the interval period that ARP request packet is sent and the retry-count after detection is failing

Command	Explanation
Interface mode	
keepalive gateway <ip-address> [{<interval-seconds> | msec <interval-millisecond >} [retry-count]] **no keepalive gateway**	Enable keepalive gateway, configure IP address of gateway, the interval period that ARP request packet is sent, and the retry-count after detection is failing, the no command disables the function.

2. Show keepalive gateway and IPv4 running status of interface

Command	Explanation
Admin and configuration mode	
show keepalive gateway [interface-name]	Show keepalive running status of the specified interface, if there is no interface is specified, show keepalive running status of all

	interfaces.
show ip interface [interface-name]	Show IPv4 running status of the specified interface, if there is no interface is specified, show IPv4 running status of all interfaces.

22.3 Keepalive Gateway Example

Keepalive gateway typical example

In above network topology, interface address of interface vlan10 is 1.1.1.1 255.255.255.0 for gateway A, interface address of interface vlan100 is 1.1.1.2 255.255.255.0 for gateway B, gateway B supports keepalive gateway function, the configuration in the following:

1. Adopt the default interval that ARP packet is sent and the retry-count after detection is failing (the default interval is 10s, the default retry-count is 5 times)

Switch(config)#interface vlan 100
Switch(config-if-vlan100)#keepalive gateway 1.1.1.1
Switch(config-if-vlan100)#exit

2. Configure the interval that ARP packet is sent and the retry-count after detection is failing manually.

Switch(config)#interface vlan 100
Switch(config-if-vlan100)#keepalive gateway 1.1.1.1 3 3
Switch(config-if-vlan100)#exit

Send ARP detection once 3 seconds to detect whether gateway A is reachable, after 3 times detection is failing, gateway A is considered to be unreachable.

22.4 Kepalive Gteway Troubleshooting

If there is any problem happens when using keepalive gateway function, please check whether the problem is caused by the following reasons:

Make sure the device is layer 3 switch, layer 2 switch does not support keepalive gateway

The detection method is used to point-to-point topology mode only

Detect IPv4 accessibility by the method, so the detection result only affects IPv4 traffic, other traffic such as IPv6 is not affected

Physical state of interface only controlled by physical signal

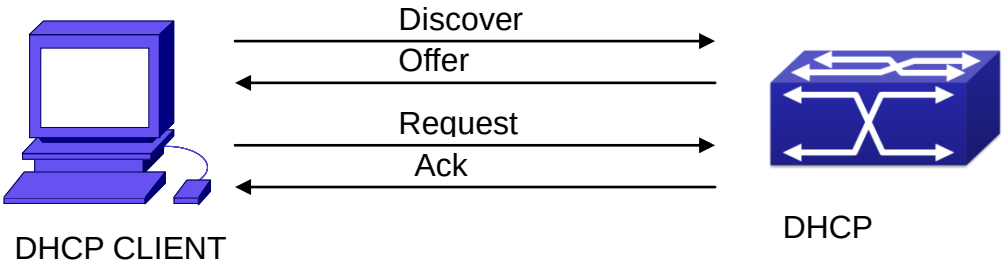
Interface can’t run IPv4 after determine gateway is not reachable, so all relative IPv4 routes are deleted and IPv4 route protocol can’t establish the neighbor on the interface

Chapter 23 DHCP Configuration

23.1 Introduction to DHCP

DHCP [RFC2131] is the acronym for Dynamic Host Configuration Protocol. It is a protocol that assigns IP address dynamically from the address pool as well as other network configuration parameters such as default gateway, DNS server, and default route and host image file position within the network. DHCP is the enhanced version of BOOTP. It is a mainstream technology that can not only provide boot information for diskless workstations, but can also release the administrators from manual recording of IP allocation and reduce user effort and cost on configuration. Another benefit of DHCP is it can partially ease the pressure on IP demands, when the user of an IP leaves the network that IP can be assigned to another user.

DHCP is a client-server protocol, the DHCP client requests the network address and configuration parameters from the DHCP server; the server provides the network address and configuration parameters for the clients; if DHCP server and clients are located in different subnets, DHCP relay is required for DHCP packets to be transferred between the DHCP client and DHCP server. The implementation of DHCP is shown below:



DHCP protocol interaction

Explanation:

DHCP client broadcasts DHCPDISCOVER packets in the local subnet.

On receiving the DHCPDISCOVER packet, DHCP server sends a DHCPOFFER packet along with IP address and other network parameters to the DHCP client.

DHCP client broadcast DHCPREQUEST packet with the information for the DHCP server it selected after selecting from the DHCPOFFER packets.

The DHCP server selected by the client sends a DHCPACK packet and the client gets an IP address and other network configuration parameters.

The above four steps finish a Dynamic host configuration assignment process. However, if the DHCP server and the DHCP client are not in the same network, the server will not receive the DHCP broadcast packets sent by the client, therefore no DHCP packets will be sent to the client by the server. In this case, a DHCP relay is required to forward such DHCP packets so that the DHCP packets exchange can be completed between the DHCP client and server.

Switch can act as both a DHCP server and a DHCP relay. DHCP server supports not only dynamic IP address assignment, but also manual IP address binding (i.e. specify a specific IP address to a specified MAC address or specified device ID over a long period. The differences and relations between dynamic IP address allocation and manual IP address binding are: 1) IP address obtained dynamically can be different every time; manually bound IP address will be the same all the time. 2) The lease period of IP address obtained dynamically is the same as the lease period of the address pool, and is limited; the lease of manually bound IP address is theoretically endless. 3) Dynamically allocated address cannot be bound manually. 4) Dynamic DHCP address pool can inherit the network configuration parameters of the dynamic DHCP address pool of the related segment.

23.2 DHCP Server Configuration

DHCP Sever Configuration Task List:
Enable/Disable DHCP service
Configure DHCP Address pool
Create/Delete DHCP Address pool
Configure DHCP address pool parameters
Configure manual DHCP address pool parameters
Enable logging for address conflicts

1. Enable/Disable DHCP service

Command	Explanation
Global Mode	
service dhcp **no service dhcp**	Enable DHCP server. The no command disables DHCP server.

2. Configure DHCP Address pool
(1) Create/Delete DHCP Address pool

Command	Explanation
Global Mode	
ip dhcp pool <*name*> **no ip dhcp pool <*name*>**	Configure DHCP Address pool. The no operation cancels the DHCP Address pool.

(2) Configure DHCP address pool parameters

Command	Explanation
DHCP Address Pool Mode	
network-address <*network-number*> [mask | prefix-length]	Configure the address scope that can be allocated to the address pool. The no

no network-address	operation of this command cancels the allocation address pool.
default-router [<address1>[<address2>[…<address8>]]] **no default-router**	Configure default gateway for DHCP clients. The no operation cancels the default gateway.
dns-server [<address1>[<address2>[…<address8>]]] **no dns-server**	Configure DNS server for DHCP clients. The no command deletes DNS server configuration.
domain-name <*domain*> **no domain-name**	Configure Domain name for DHCP clients; the "**no domain-name**" command deletes the domain name.
netbios-name-server [<address1>[<address2>[…<address8>]]] **no netbios-name-server**	Configure the address for WINS server. The no operation cancels the address for server.
netbios-node-type {b-node|h-node|m-node|p-node|<*type-number*>} **no netbios-node-type**	Configure node type for DHCP clients. The no operation cancels the node type for DHCP clients.
bootfile <filename> **no bootfile**	Configure the file to be imported for DHCP clients on boot up. The no command cancels this operation.
next-server [<address1>[<address2>[…<address8>]]] **no next-server [<address1>[<address2>[…<address8>]]]**	Configure the address of the server hosting file for importing. The no command deletes the address of the server hosting file for importing.
option <*code*> {ascii <*string*> | hex <*hex*> | ipaddress <*ipaddress*>} **no option <*code*>**	Configure the network parameter specified by the option code. The no command deletes the network parameter specified by the option code.
lease { days [hours][minutes] | infinite } **no lease**	Configure the lease period allocated to addresses in the address pool. The no command deletes the lease period allocated to addresses in the address pool.
Global Mode	
ip dhcp excluded-address <*low-address*> [<*high-address*>] **no ip dhcp excluded-address <*low-address*> [<*high-address*>]**	Exclude the addresses in the address pool that are not for dynamic allocation.

(3) Configure manual DHCP address pool parameters

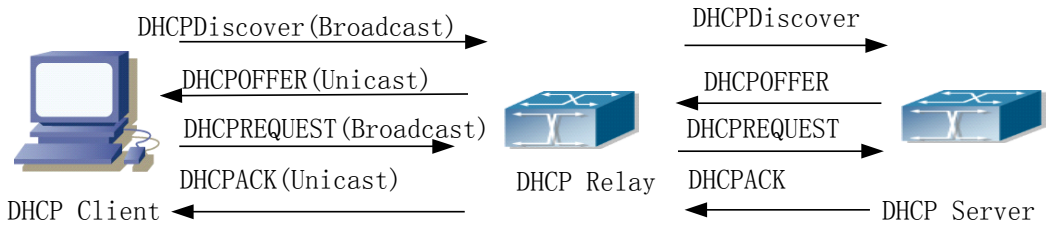
Command	Explanation
DHCP Address Pool Mode	
hardware-address <*hardware-address*> [{Ethernet | IEEE802|<*type-number*>}] **no hardware-address**	Specify/delete the hardware address when assigning address manually.
host <address> [<mask> | <prefix-length>] **no host**	Specify/delete the IP address to be assigned to the specified client when binding address manually.
client-identifier <unique-identifier> **no client-identifier**	Specify/delete the unique ID of the user when binding address manually.

3. Enable logging for address conflicts

Command	Explanation
Global Mode	
ip dhcp conflict logging **no ip dhcp conflict logging**	Enable/disable logging for DHCP address to detect address conflicts.
Admin Mode	
clear ip dhcp conflict <*address* | all >	Delete a single address conflict record or all conflict records.

23.3 DHCP Relay Configuration

When the DHCP client and server are in different segments, DHCP relay is required to transfer DHCP packets. Adding a DHCP relay makes it unnecessary to configure a DHCP server for each segment, one DHCP server can provide the network configuration parameter for clients from multiple segments, which is not only cost-effective but also management-effective.



DHCP relay

As shown in the above figure, the DHCP client and the DHCP server are in different networks, the DHCP client performs the four DHCP steps as usual yet DHCP relay is added to the process.

The client broadcasts a DHCPDISCOVER packet, and DHCP relay inserts its own IP address

to the relay agent field in the DHCPDISCOVER packet on receiving the packet, and forwards the packet to the specified DHCP server (for DHCP frame format, please refer to RFC2131).
On the receiving the DHCPDISCOVER packets forwarded by DHCP relay, the DHCP server sends the DHCPOFFER packet via DHCP relay to the DHCP client.
DHCP client chooses a DHCP server and broadcasts a DHCPREQUEST packet, DHCP relay forwards the packet to the DHCP server after processing.
On receiving DHCPREQUEST, the DHCP server responds with a DHCPACK packet via DHCP relay to the DHCP client.
DHCP Relay Configuration Task List:
1. Enable DHCP relay.
2. Configure DHCP relay to forward DHCP broadcast packet.

1. Enable DHCP relay.

Command	Explanation
Global Mode	
service dhcp **no service dhcp**	DHCP server and DHCP relay is enabled as the DHCP service is enabled.

2. Configure DHCP relay to forward DHCP broadcast packet.

Command	Explanation
Global Mode	
ip forward-protocol udp bootps **no ip forward-protocol udp bootps**	The UDP port 67 is used for DHCP broadcast packet forwarding.
Interface Configuration Mode	
ip helper-address <*ipaddress*> **no ip helper-address <*ipaddress*>**	Set the destination IP address for DHCP relay forwarding; the "**no ip helper-address <*ipaddress*>**"command cancels the setting.

23.4 DHCP Configuration Examples

Scenario 1:
Too save configuration efforts of network administrators and users, a company is using switch as a DHCP server. The Admin VLAN IP address is 10.16.1.2/16. The local area network for the company is divided into network A and B according to the office locations. The network configurations for location A and B are shown below.

PoolA(network 10.16.1.0)		**PoolB(network 10.16.2.0)**	
Device	IP address	**Device**	IP address
Default gateway	10.16.1.200	**Default gateway**	10.16.1.200

	10.16.1.201		10.16.1.201
DNS server	10.16.1.202	**DNS server**	10.16.1.202
WINS server	10.16.1.209	**WWW server**	10.16.1.209
WINS node type	H-node		
Lease	3 days	**Lease**	1day

In location A, a machine with MAC address 00-03-22-23-dc-ab is assigned with a fixed IP address of 10.16.1.210 and named as “management”.

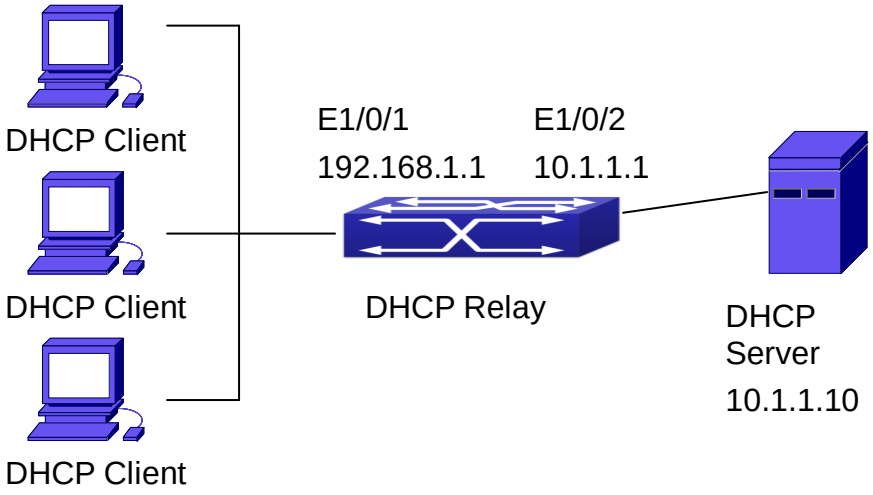
```
Switch(config)#service dhcp
Switch(config)#interface vlan 1
Switch(Config-Vlan-1)#ip address 10.16.1.2 255.255.0.0
Switch(Config-Vlan-1)#exit
Switch(config)#ip dhcp pool A
Switch(dhcp-A-config)#network 10.16.1.0 24
Switch(dhcp-A-config)#lease 3
Switch(dhcp-A-config)#default-route 10.16.1.200 10.16.1.201
Switch(dhcp-A-config)#dns-server 10.16.1.202
Switch(dhcp-A-config)#netbios-name-server 10.16.1.209
Switch(dhcp-A-config)#netbios-node-type H-node
Switch(dhcp-A-config)#exit
Switch(config)#ip dhcp excluded-address 10.16.1.200 10.16.1.201
Switch(config)#ip dhcp pool B
Switch(dhcp-B-config)#network 10.16.2.0 24
Switch(dhcp-B-config)#lease 1
Switch(dhcp-B-config)#default-route 10.16.2.200 10.16.2.201
Switch(dhcp-B-config)#dns-server 10.16.2.202
Switch(dhcp-B-config)#option 72 ip 10.16.2.209
Switch(dhcp-config)#exit
Switch(config)#ip dhcp excluded-address 10.16.2.200 10.16.2.201
Switch(config)#ip dhcp pool A1
Switch(dhcp-A1-config)#host 10.16.1.210
Switch(dhcp-A1-config)#hardware-address 00-03-22-23-dc-ab
Switch(dhcp-A1-config)#exit
```

Usage Guide: When a DHCP/BOOTP client is connected to a VLAN1 port of the switch, the client can only get its address from 10.16.1.0/24 instead of 10.16.2.0/24. This is because the broadcast packet from the client will be requesting the IP address in the same segment of the VLAN interface after VLAN interface forwarding, and the VLAN interface IP address is 10.16.1.2/24, therefore the IP address assigned to the client will belong to 10.16.1.0/24.

If the DHCP/BOOTP client wants to have an address in 10.16.2.0/24, the gateway forwarding broadcast packets of the client must belong to 10.16.2.0/24. The connectivity between the

client gateway and the switch must be ensured for the client to get an IP address from the 10.16.2.0/24 address pool.

Scenario 2:



DHCP Relay Configuration

As shown in the above figure, route switch is configured as a DHCP relay. The DHCP server address is 10.1.1.10, the configuration steps is as follows:

```
Switch(config)#service dhcp
Switch(config)#interface vlan 1
Switch(Config-if-Vlan1)#ip address 192.168.1.1 255.255.255.0
Switch(Config-if-Vlan1)#exit
Switch(config)#vlan 2
Switch(Config-Vlan-2)#exit
Switch(config)#interface Ethernet 1/0/2
Switch(Config-Erthernet1/0/2)#switchport access vlan 2
Switch(Config-Erthernet1/0/2)#exit
Switch(config)#interface vlan 2
Switch(Config-if-Vlan2)#ip address 10.1.1.1 255.255.255.0
Switch(Config-if-Vlan2)#exit
Switch(config)#ip forward-protocol udp bootps
Switch(config)#interface vlan 1
Switch(Config-if-Vlan1)#ip help-address 10.1.1.10
Switch(Config-if-Vlan1)#exit
```

Note: It is recommended to use the combination of command **ip forward-protocol udp *<port>*** and **ip helper-address *<ipaddress>***. **ip help-address** can only be configured for ports on layer 3 and cannot be configured on layer 2 ports directly.

23.5 DHCP Troubleshooting

If the DHCP clients cannot obtain IP addresses and other network parameters, the following procedures can be followed when DHCP client hardware and cables have been verified ok.

Verify the DHCP server is running, start the related DHCP server if not running. If the DHCP clients and servers are not in the same physical network, verify the router responsible for DHCP packet forwarding has DHCP relay function. If DHCP relay is not available for the intermediate router, it is recommended to replace the router or upgrade its software to one that has a DHCP relay function.

In such case, DHCP server should be examined for an address pool that is in the same segment of the switch VLAN, such a pool should be added if not present, and (This does not indicate switch cannot assign IP address for different segments, see solution 2 for details.)

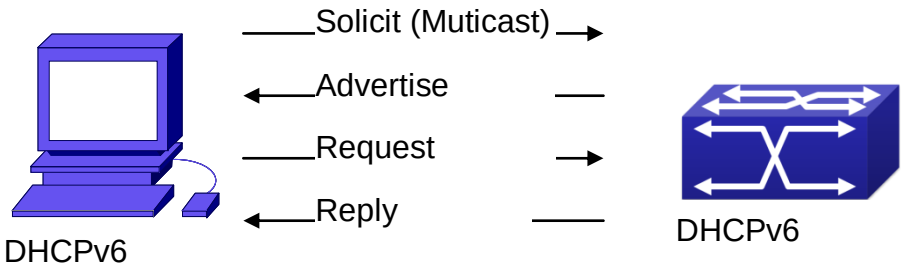
In DHCP service, pools for dynamic IP allocation and manual binding are conflicting, i.e., if command “**network-address**” and “**host**” are run for a pool, only one of them will take effect; furthermore, in manual binding, only one IP-MAC binding can be configured in one pool. If multiple bindings are required, multiple manual pools can be created and IP-MAC bindings set for each pool. New configuration in the same pool overwrites the previous configuration.

Chapter 24 DHCPv6 Configuration

24.1 Introduction to DHCPv6

DHCPv6 [RFC3315] is the IPv6 version for Dynamic Host Configuration Protocol (DHCP). It is a protocol that assigns IPv6 address as well as other network configuration parameters such as DNS address, and domain name to DHCPv6 client, DHCPv6 is a conditional auto address configuration protocol relative to IPv6. In the conditional address configuration process, DHCPv6 server assigns a complete IPv6 address to client, and provides DNS address, domain name and other configuration information, maybe the DHCPv6 packet can transmit through relay delegation, at last the binding of IPv6 address and client can be recorded by DHCPv6 server, all that can enhance the management of network; DHCPv6 server can also provide non state DHCPv6 service, that is only assigns DNS address and domain name and other configuration information but not assigns IPv6 address, it can solve the bug of IPv6 auto address configuration in non state; DHCPv6 can provide extend function of DHCPv6 prefix delegation, upstream route can assign address prefix to downstream route automatically, that achieve the IPv6 address auto assignment in levels of network environment, and resolved the problem of ISP and IPv6 network dispose.

There are three entities in the DHCPv6 protocol – the client, the relay and the server. The DHCPv6 protocol is based on the UDP protocol. The DHCPv6 client sends request messages to the DHCP server or DHCP relay with the destination port as 547, and the DHCPv6 server and relay send replying messages with the destination port as 546. The DHCPv6 client sends solicit or request messages with the multicast address – ff02::1:2 for DHCP relay and server.



DHCPv6 negotiation

When a DHCPv6 client tries to request an IPv6 address and other configurations from the DHCPv6 server, the client has to find the location of the DHCP server, and then request configurations from the DHCP server.

In the time of located server, the DHCP client tries to find a DHCPv6 server by broadcasting a SOLICIT packet to all the DHCP delay delegation and server with broadcast address as FF02::1:2.

Any DHCP server which receives the request, will reply the client with an ADVERTISE

message, which includes the identity of the server –DUID, and its priority.
It is possible that the client receives multiple ADVERTISE messages. The client should select one and reply it with a REQUEST message to request the address which is advertised in the ADVERTISE message.
The selected DHCPv6 server then confirms the client about the IPv6 address and any other configuration with the REPLY message.
The above four steps finish a Dynamic host configuration assignment process. However, if the DHCPv6 server and the DHCPv6 client are not in the same network, the server will not receive the DHCPv6 broadcast packets sent by the client, therefore no DHCPv6 packets will be sent to the client by the server. In this case, a DHCPv6 relay is required to forward such DHCPv6 packets so that the DHCPv6 packets exchange can be completed between the DHCPv6 client and server.
At the time this manual is written, DHCPv6 server, relay and prefix delegation client have been implemented on the switch. When the DHCPv6 relay receives any messages from the DHCPv6 client, it will encapsulate the request in a Relay-forward packet and deliver it to the next DHCPv6 relay or the DHCPv6 server. The DHCPv6 messages coming from the server will be encapsulated as relay reply packets to the DHCPv6 relay. The relay then removes the encapsulation and delivers it the DHCPv6 client or the next DHCPv6 relay in the network.
For DHCPv6 prefix delegation where DHCPv6 server is configured on the PE router and DHCPv6 client it configured on the CPE router, the CPE router is able to send address prefix allocation request to the PE router and get a pre-configured address prefix, but not set the address prefix manually. The protocol negotiation between the client and the prefix delegation client is quite similar to that when getting a DHCPv6 address. Then the CPE router divides the allocated prefix – whose length should be less than 64 characters, into 64 subnets. The divided address prefix will be advertised through routing advertisement messages (RA) to the host directly connected to the client.

24.2 DHCPv6 Server Configuration

DHCPv6 server configuration task list as below:
To enable/disable DHCPv6 service
To configure DHCPv6 address pool
(1) To achieve/delete DHCPv6 address pool
(2) To configure parameter of DHCPv6 address pool
To enable DHCPv6 server function on port

1. To enable/disable DHCPv6 service

Command	Explanation
Global Mode	
service dhcpv6 **no service dhcpv6**	To enable DHCPv6 service.

2. To configure DHCPv6 address pool

(1) To achieve/delete DHCPv6 address pool

Command	Explanation
Global Mode	
ipv6 dhcp pool <*poolname*> **no ipv6 dhcp pool <*poolname*>**	To configure DHCPv6 address pool.

(2) To configure parameter of DHCPv6 address pool

Command	Explanation
DHCPv6 address pool Configuration Mode	
network-address <ipv6-pool-start-address> {<ipv6-pool-end-address> | <prefix-length>} [eui-64] **no network-address**	To configure the range of IPv6 address assignable of address pool.
dns-server <ipv6-address> **no dns-server <*ipv6-address*>**	To configure DNS server address for DHCPv6 client.
domain-name <domain-name> **no domain-name <*domain-name*>**	To configure DHCPv6 client domain name.
excluded-address <*ipv6-address*> **no excluded-address <*ipv6-address*>**	To exclude IPv6 address which isn't used for dynamic assignment in address pool.
lifetime {<*valid-time*> | infinity} {<*preferred-time*> | infinity} **no lifetime**	To configure valid time or preferred time of DHCPv6 address pool.

3. To enable DHCPv6 server function on port.

Command	Explanation
Interface Configuration Mode	
ipv6 dhcp server <*poolname*> [preference <*value*>] [rapid-commit] [allow-hint] **no ipv6 dhcp server <*poolname*>**	To enable DHCPv6 server function on specified port, and binding the used DHCPv6 address pool.

24.3 DHCPv6 Relay Delegation Configuration

DHCPv6 relay delegation configuration task list as below:
To enable/disable DHCPv6 service
To configure DHCPv6 relay delegation on port

1. To enable DHCPv6 service

Command	Explanation
Global Mode	
service dhcpv6 **no service dhcpv6**	To enableDHCPv6 service.

2. To configure DHCPv6 relay delegation on port

Command	Explanation
Interface Configuration Mode	
ipv6 dhcp relay destination {[<*ipv6-address*>] [interface { <*interface-name*> | vlan <*1-4096*>}]} **no ipv6 dhcp relay destination {[<*ipv6-address*>] [interface { <*interface-name*> | vlan <*1-4096*>}]}**	To specify the destination address of DHCPv6 relay transmit; The no form of this command delete the configuration.

24.4 DHCPv6 Prefix Delegation Server Configuration

DHCPv6 prefix delegation server configuration task list as below:
To enable/delete DHCPv6 service
To configure prefix delegation pool
To configure DHCPv6 address pool
(1) To achieve/delete DHCPv6 address pool
(2) To configure prefix delegation pool used by DHCPv6 address pool
(3) To configure static prefix delegation binding
(4) To configure other parameters of DHCPv6 address pool
To enable DHCPv6 prefix delegation server function on port

1. To enable/delete DHCPv6 service

Command	Explanation
Global Mode	
service dhcpv6	To enable DHCPv6 service.

no service dhcpv6	

2. To configure prefix delegation pool

Command	Explanation
Global Mode	
ipv6 local pool <poolname> <prefix|prefix-length> <assigned-length> **no ipv6 local pool <*poolname*>**	To configure prefix delegation pool.

3. To configure DHCPv6 address pool

(1) To achieve/delete DHCPv6 address pool

Command	Explanation
Global Mode	
ipv6 dhcp pool <*poolname*> **no ipv6 dhcp pool <*poolname*>**	To configure DHCPv6 address pool.

(2) To configure prefix delegation pool used by DHCPv6 address pool

Command	Explanation
DHCPv6 address pool Configuration Mode	
prefix-delegation pool <*poolname*> **[lifetime {<*valid-time*> | infinity}** **{<*preferred-time*> | infinity}]** **no prefix-delegation pool <*poolname*>**	To specify prefix delegation pool used by DHCPv6 address pool, and assign usable prefix to client.

(3) To configure static prefix delegation binding

Command	Explanation
DHCPv6 address pool Configuration Mode	
prefix-delegation <*ipv6-prefix/prefix-length*> <*client-DUID*> [iaid <*iaid*>] **[lifetime {<valid-time> | infinity}** **{<*preferred-time*> | infinity}]** **no prefix-delegation <ipv6-prefix/prefix-length> <client-DUID> [iaid <iaid>]**	To specify IPv6 prefix and any prefix required static binding by client.

(4) To configure other parameter of DHCPv6 address pool

Command	Explanation
DHCPv6 address pool Configuration Mode	
dns-server <ipv6-address> **no dns-server <*ipv6-address*>**	To configure DNS server address for DHCPv6 client.
domain-name <domain-name> **no domain-name <*domain-name*>**	To configure domain name for DHCPv6 client.

4. To enable DHCPv6 prefix delegation server function on port

Command	Explanation
Interface Configuration Mode	
ipv6 dhcp server <*poolname*> [preference <*value*>] [rapid-commit] [allow-hint] **no ipv6 dhcp server <*poolname*>**	To enable DHCPv6 server function on specified port, and binding used DHCPv6 address pool.

24.5 DHCPv6 Prefix Delegation Client Configuration

DHCPv6 prefix delegation client configuration task list as below:
To enable/disable DHCPv6 service
To enable DHCPv6 prefix delegation client function on port

1. To enable/disable DHCPv6 service

Command	Explanation
Global Mode	
service dhcpv6 **no service dhcpv6**	To enable DHCPv6 service.

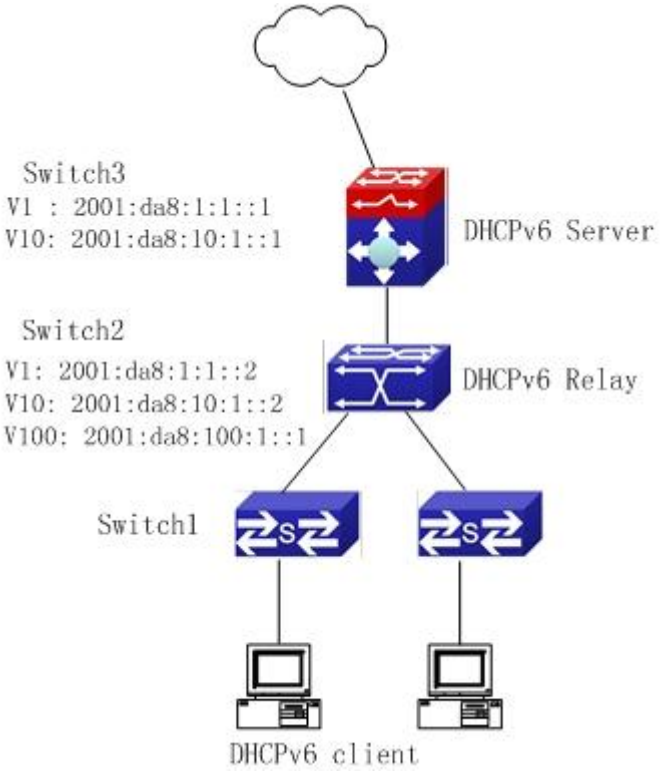
2. To enable DHCPv6 prefix delegation client function on port

Command	Explanation
Interface Configuration Mode	
ipv6 dhcp client pd <*prefix-name*> [rapid-commit] **no ipv6 dhcp client pd**	To enable client prefix delegation request function on specified port, and the prefix obtained associate with universal prefix configured.

24.6 DHCPv6 Configuration Examples

Example1:
When deploying IPv6 networking, QTECH series switches can be configured as DHCPv6 server in order to manage the allocation of IPv6 addresses. Both the state and the stateless DHCPv6 are supported.
Topology:
The access layer use Switch1 switch to connect users of dormitory buildings; Switch2 is configured as DHCPv6 relay delegation in primary aggregation layer ; Switch3 is configured as

DHCPv6 server in secondary aggregation layer, and connected with backbone network or higher aggregation layers; The Windows Vista which be provided with DHCPv6 client must load on PC.



Usage guide:

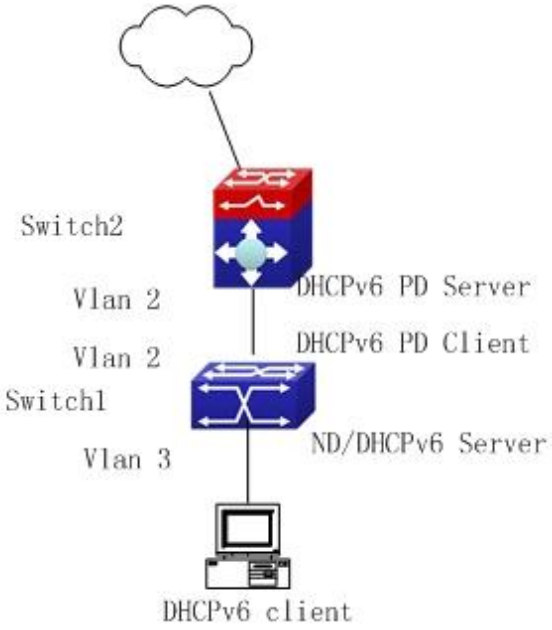
Switch3 configuration:

```
Switch3>enable
Switch3#config
Switch3(config)#service dhcpv6
Switch3(config)#ipv6 dhcp pool EastDormPool
Switch3(dhcpv6-EastDormPool-config)#network-address 2001:da8:100:1::1 2001:da8:100:1::100
Switch3(dhcpv6-EastDormPool-config)#excluded-address 2001:da8:100:1::1
Switch3(dhcpv6-EastDormPool-config)#dns-server 2001:da8::20
Switch3(dhcpv6-EastDormPool-config)#dns-server 2001:da8::21
Switch3(dhcpv6-EastDormPool-config)#domain-name dhcpv6.com
Switch3(dhcpv6-EastDormPool-config)#lifetime 1000 600
Switch3(dhcpv6-EastDormPool-config)#exit
Switch3(config)#interface vlan 1
Switch3(Config-if-Vlan1)#ipv6 address 2001:da8:1:1::1/64
Switch3(Config-if-Vlan1)#exit
Switch3(config)#interface vlan 10
Switch3(Config-if-Vlan10)#ipv6 address 2001:da8:10:1::1/64
Switch3(Config-if-Vlan10)#ipv6 dhcp server EastDormPool preference 80
```

Switch3(Config-if-Vlan10)#exit
Switch3(config)#

Switch2 configuration:
Switch2>enable
Switch2#config
Switch2(config)#service dhcpv6
Switch2(config)#interface vlan 1
Switch2(Config-if-Vlan1)#ipv6 address 2001:da8:1:1::2/64
Switch2(Config-if-Vlan1)#exit
Switch2(config)#interface vlan 10
Switch2(Config-if-Vlan10)#ipv6 address 2001:da8:10:1::2/64
Switch2(Config-if-Vlan10)#exit
Switch2(config)#interface vlan 100
Switch2(Config-if-Vlan100)#ipv6 address 2001:da8:100:1::1/64
Switch2(Config-if-Vlan100)#no ipv6 nd suppress-ra
Switch2(Config-if-Vlan100)#ipv6 nd managed-config-flag
Switch2(Config-if-Vlan100)#ipv6 nd other-config-flag
Switch2(Config-if-Vlan100)#ipv6 dhcp relay destination 2001:da8:10:1::1
Switch2(Config-if-Vlan100)#exit
Switch2(config)#
Example2:
When the network operator is deploying IPv6 networks, network automatically configuration can be achieved through the prefix delegation allocation of IPv6 addresses, in stead of configuring manually for each switch:
To configure the switching or routing device which is connected to the client switch as DHCPv6 prefix delegation server, that is to setup a local database for the relationship between the allocated prefix and the DUID of the client switch.
To configure the switch as the prefix delegation client, and make the client switch to get IPv6 address prefix from the prefix delegation server, through a process which is much like the process of DHCPv6 address allocation.
The edge devices which receive the address prefix, send routing advertisement - RA messages, to the client hosts about the address prefix through the interface which is connected to the hosts, then the hosts get an valid IPv6 address through stateless auto configuration, while at the same time, the stateless DHCPv6 server will be configured for the interface, in order to provide the DHCPv6 client with information such as DNS, and domain name, etc.
Network Topology:
The edge switch is a Switch1 switch. The interface connected to the trunk switch which is

Switch2, is configured as the prefix delegation client. The interfaces connected to hosts, are configured as stateless DHCPv6 servers to provide the hosts with stateless information such as DNS and domain names, also routing advertisement of stateless address allocation is enabled for the host interfaces; On Switch2, the prefix delegation server is configured, and routing advertisement of state address allocation is enabled; On the host side, DHCPv6 client capable operating system such Windows Vista should be installed.



Usage guide:

Switch2 configuration

Switch2>enable

Switch2#config

Switch2(config)#interface vlan 2

Switch2(Config-if-Vlan2)#ipv6 address 2001:da8:1100::1/64

Switch2(Config-if-Vlan2)#exit

Switch2(config)#service dhcpv6

Switch2(config)#ipv6 local pool client-prefix-pool 2001:da8:1800::/40 48

Switch2(config)#ipv6 dhcp pool dhcp-pool

Switch2(dhcpv6-dhcp-pool-config)#prefix-delegation pool client-prefix-pool 1800 600

Switch2(dhcpv6-dhcp-pool-config)#exit

Switch2(config)#interface vlan 2

Switch2(Config-if-Vlan2)#ipv6 dhcp server dhcp-pool

Switch2(Config-if-Vlan2)#exit

Switch1 configuration

Switch1>enable

Switch1#config

Switch1(config)#service dhcpv6

```
Switch1(config)#interface vlan 2
Switch1(Config-if-Vlan2)#ipv6 dhcp client pd prefix-from-provider
Switch1(Config-if-Vlan2)#exit
Switch1(config)#interface vlan 3
Switch1(Config-if-Vlan3)#ipv6 address prefix-from-provider 0:0:0:1::1/64
Switch1(Config-if-Vlan3)#exit
Switch1(config)#ipv6 dhcp pool foo
Switch1(dhcpv6-foo-config)#dns-server 2001:4::1
Switch1(dhcpv6-foo-config)#domain-name www.ipv6.org
Switch1(dhcpv6-foo-config)#exit
Switch1(config)#interface vlan 3
Switch1(Config-if-Vlan3)#ipv6 dhcp server foo
Switch1(Config-if-Vlan3)#ipv6 nd other-config-flag
Switch1(Config-if-Vlan3)#no ipv6 nd suppress-ra
Switch1(Config-if-Vlan3)#exit
```

24.7 DHCPv6 Troubleshooting

If the DHCPv6 clients cannot obtain IPv6 addresses and other network parameters, the following procedures can be followed when DHCPv6 client hardware and cables have been verified ok:

Verify the DHCPv6 server is running, start the related DHCP v6 server function if not running;

If the DHCPv6 clients and servers are not in the same physical network, verify the router responsible for DHCPv6 packet forwarding has DHCPv6 relay function. If DHCPv6 relay is not available for the intermediate router, it is recommended to replace the router or upgrade its software to one that has a DHCPv6 relay function;

Sometimes hosts are connected to the DHCPv6 enabled switches, but can not get IPv6 addresses. In this situation, it should be checked first whether the ports which the hosts are connected to, are connected with the port which the DHCPv6 server is connected to. If connected directly, it should be checked then whether the IPv6 address pool of the VLAN which the port belongs to, is in the same subnet with the address pool configure in the DHCPv6 server; If not connected directly, and any layer three DHCPv6 relay is configured between the hosts and the DHCPv6 server, it should be checked first whether an valid IPv6 address has been configured for the switch interface which the hosts are connected to. If not configured, configure an valid IPv6 address. If configured, it should be checked whether the configured IPv6 address is in the same subnet with the DHCPv6 server. If not, please add it to the address pool.

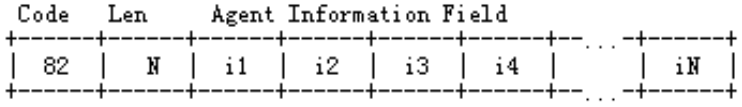
Chapter 25 DHCP option 82 Configuration

25.1 Introduction to DHCP option 82

DHCP option 82 is the Relay Agent Information Option, its option code is 82. DHCP option 82 is aimed at strengthening the security of DHCP servers and improving the IP address configuration policy. The Relay Agent adds option 82 (including the client's physical access port, the access device ID and other information), to the DHCP request message from the client then forwards the message to DHCP server. When the DHCP server which supports the option 82 function receives the message, it will allocate an IP address and other configuration information for the client according to preconfigured policies and the option 82 information in the message. At the same time, DHCP server can identify all the possible DHCP attack messages according to the information in option 82 and defend against them. DHCP Relay Agent will peel the option 82 from the reply messages it receives, and forward the reply message to the specified port of the network access device, according to the physical port information in the option. The application of DHCP option 82 is transparent for the client.

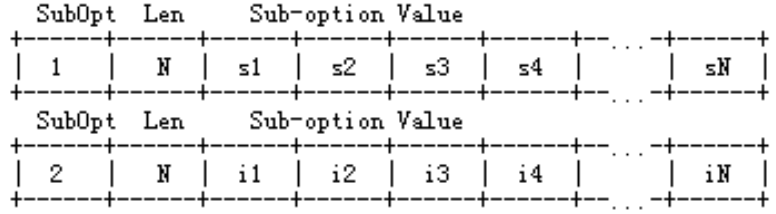
25.1.1 DHCP option 82 Message Structure

A DHCP message can have several option segments; option 82 is one of them. It has to be placed after other options but before option 255. The following is its format:



Code: represents the sequence number of the relay agent information option, the option 82 is called so because RFC3046 is defined as 82.
Len: the number of bytes in Agent Information Field, not including the two bytes in Code segment and Len segment.

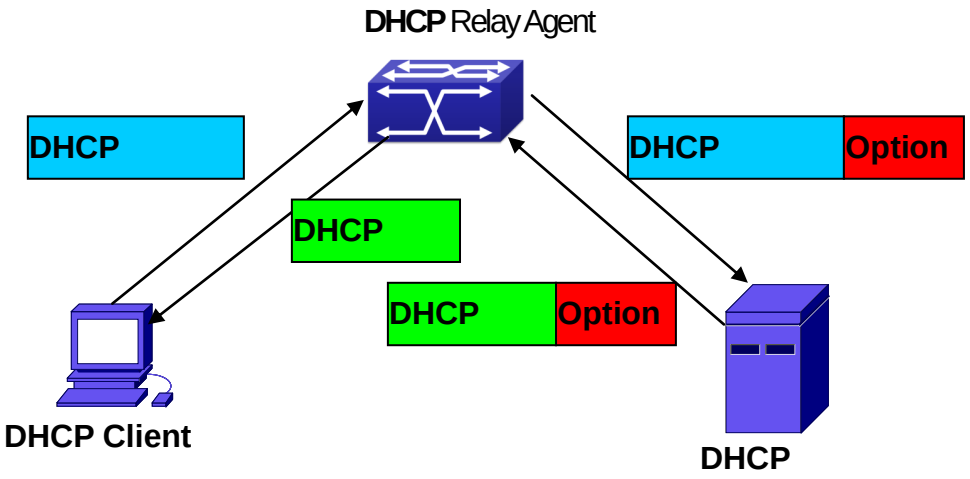
Option 82 can have several sub-options, and need at least one sub-option. RFC3046 defines the following two sub-options, whose formats are showed as follows:



SubOpt: the sequence number of sub-option, the sequence number of Circuit ID sub-option is 1, the sequence number of Remote ID sub-option is 2.

Len: the number of bytes in Sub-option Value, not including the two bytes in SubOpt segment and Len segment.

25.1.2 Option 82 Working Mechanism



DHCP option 82 flow chart

If the DHCP Relay Agent supports option 82, the DHCP client should go through the following four steps to get its IP address from the DHCP server: discover, offer, select and acknowledge. The DHCP protocol follows the procedure below:

1) DHCP client sends a request broadcast message while initializing. This request message does not have option 82.

2) DHCP Relay Agent will add the option 82 to the end of the request message it receives, then relay and forward the message to the DHCP server. By default, the sub-option 1 of option 82 (Circuit ID) is the interface information of the switch connected to the DHCP client (VLAN name and physical port name), but the users can configure the Circuit ID as they wish. The sub-option 2 of option 82(Remote ID) is the MAC address of the DHCP relay device.

3) After receiving the DHCP request message, the DHCP server will allocate IP address and other information for the client according to the information and preconfigured policy in the option segment of the message. Then it will forward the reply message with DHCP configuration information and option 82 information to DHCP Relay Agent.

4) DHCP Relay Agent will peel the option 82 information from the replay message sent by DHCP server, and then forward the message with DHCP configuration information to the DHCP client.

25.2 DHCP option 82 Configuration Task List

Enabling the DHCP option 82 of the Relay Agent
Configure the DHCP option 82 attributes of the interface
Enable the DHCP option 82 of server

Configure DHCP option 82 default format of Relay Agent
Configure delimiter
Configure creation method of option82
Diagnose and maintain DHCP option 82

1. Enabling the DHCP option 82 of the Relay Agent.

Command	Explanation
Global mode	
ip dhcp relay information option **no ip dhcp relay information option**	Set this command to enable the option 82 function of the switch Relay Agent. The “no ip dhcp relay information option” is used to disable the option 82 function of the switch Relay Agent.

2. Configure the DHCP option 82 attributes of the interface

Command	Explanation
Interface configuration mode	
ip dhcp relay information policy {drop | keep | replace} **no ip dhcp relay information policy**	This command is used to set the retransmitting policy of the system for the received DHCP request message which contains option 82. The drop mode means that if the message has option82, then the system will drop it without processing; keep mode means that the system will keep the original option 82 segment in the message, and forward it to the server to process; replace mode means that the system will replace the option 82 segment in the existing message with its own option 82, and forward the message to the server to process. The “no ip dhcp relay information policy” will set the retransmitting policy of the option 82 DCHP message as “replace”.
ip dhcp relay information option subscriber-id {standard | <*circuit-id*>} **no ip dhcp relay information option subscriber-id**	This command is used to set the format of option 82 sub-option1(Circuit ID option) added to the DHCP request messages from interface, standard means the standard VLAN name and physical port name format,

	like"Vlan2+Ethernet1/0/12",<circuit-id> is the circuit-id contents of option 82 specified by users, which is a string no longer than 64characters. The" **no ip dhcp relay information option subscriber-id**" command will set the format of added option 82 sub-option1 (Circuit ID option) as standard format.
Global Mode	
ip dhcp relay information option remote-id {standard | *<remote-id>*} **no ip dhcp relay information option remote-id**	Set the suboption2 (remote ID option) content of option 82 added by DHCP request packets (They are received by the interface). The no command sets the additive suboption2 (remote ID option) format of option 82 as standard.

3. Enable the DHCP option 82 of server.

Command	Explanation
Global mode	
ip dhcp server relay information enable **no ip dhcp server relay information enable**	This command is used to enable the switch DHCP server to identify option82. The "**no ip dhcp server relay information enable**" command will make the server ignore the option 82.

4. Configure DHCP option 82 default format of Relay Agent

Command	Explanation
Global mode	
ip dhcp relay information option subscriber-id format {hex | acsii | vs-hp}	Set subscriber-id format of Relay Agent option82.
ip dhcp relay information option remote-id format {default | vs-hp}	Set remote-id format of Relay Agent option82.

5. Configure delimiter

Command	Explanation
Global mode	
ip dhcp relay information option delimiter [colon | dot | slash | space]	Set the delimiter of each parameter for suboption of option82 in global mode, no

no ip dhcp relay information option delimiter	command restores the delimiter as slash.

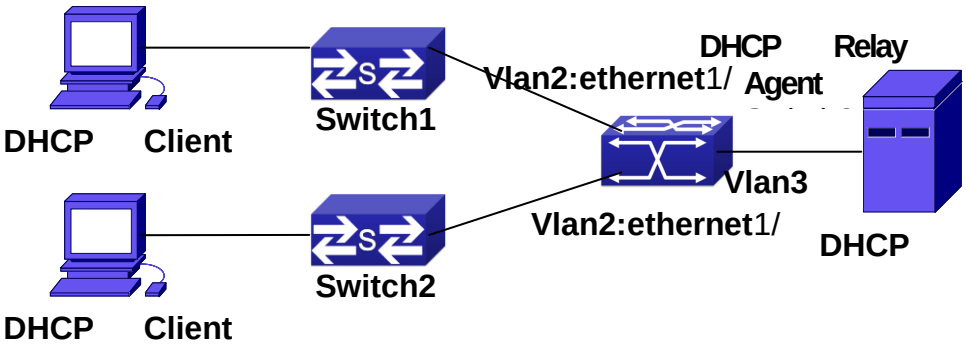
6. Configure creation method of option82

Command	Explanation
Global mode	
ip dhcp relay information option self-defined remote-id {hostname | mac | string WORD} **no ip dhcp relay information option self-defined remote-id**	Set creation method for option82, users can define the parameters of remote-id suboption by themselves
ip dhcp relay information option self-defined remote-id format [ascii | hex]	Set self-defined format of remote-id for relay option82.
ip dhcp relay information option self-defined subscriber-id {vlan | port | id (switch-id (mac | hostname)| remote-mac)| string WORD } **no ip dhcp relay information option self-defined subscriber-id**	Set creation method for option82, users can define the parameters of circute-id suboption by themselves
ip dhcp relay information option self-defined subscriber-id format [ascii | hex]	Set self-defined format of circuit-id for relay option82.

7. Diagnose and maintain DHCP option 82

Command	Explanation
Admin mode	
show ip dhcp relay information option	This command will display the state information of the DHCP option 82 in the system, including option82 enabling switch, the interface retransmitting policy, the circuit ID mode and the DHCP server option82 enabling switch.
debug ip dhcp relay packet	This command is used to display the information of data packets processing in DHCP Relay Agent, including the “add” and “peel” action of option 82.

25.3 DHCP option 82 Application Examples



A DHCP option 82 typical application example

In the above example, layer 2 switches Switch1 and Switch2 are both connected to layer 3 switch Switch3, Switch 3 will transmit the request message from DHCP client to DHCP serer as DHCP Relay Agent. It will also transmit the reply message from the server to DHCP client to finish the DHCP protocol procedure. If the DHCP option 82 is disabled, DHCP server cannot distinguish that whether the DHCP client is from the network connected to Switch1 or Switch2. So, all the PC terminals connected to Switch1 and Switch2 will get addresses from the public address pool of the DHCP server. After the DHCP option 82 function is enabled, since the Switch3 appends the port information of accessing Switch3 to the request message from the client, the server can tell that whether the client is from the network of Swich1 or Swich2, and thus can allocate separate address spaces for the two networks, to simplify the management of networks.

The following is the configuration of Switch3(MAC address is 00:1f:ce:02:33:01):
Switch3(Config)#service dhcp
Switch3(Config)#ip dhcp relay information option
Switch3(Config)#ip forward-protocol udp bootps
Switch3(Config)#interface vlan 3
Switch3(Config-if-vlan3)#ip address 192.168.10.222 255.255.255.0
Switch3(Config-if-vlan2)#ip address 192.168.102.2 255.255.255.0
Switch3(Config-if-vlan2)#ip helper 192.168.10.88

Linux ISC DHCP Server supports option 82, its configuration file /etc/dhcpd.con is ddns-update-style interim;
ignore client-updates;

class "Switch3Vlan2Class1" {
match if option agent.circuit-id = "Vlan2+Ethernet1/0/2" and option agent.remote-id=00:1f:ce:02:33:01;

```
}

class "Switch3Vlan2Class2" {
match if option agent.circuit-id = "Vlan2+Ethernet1/0/3" and option agent.remote-
id=00:1f:ce:02:33:01;
}

subnet 192.168.102.0 netmask 255.255.255.0 {
option routers 192.168.102.2;
option subnet-mask 255.255.255.0;
option domain-name "example.com.cn";
option domain-name-servers 192.168.10.3;
authoritative;

pool {
range 192.168.102.21 192.168.102.50;
default-lease-time 86400; #24 Hours
max-lease-time 172800; #48 Hours
allow members of "Switch3Vlan2Class1";
}
pool {
range 192.168.102.51 192.168.102.80;
default-lease-time 43200; #12 Hours
max-lease-time 86400; #24 Hours
allow members of "Switch3Vlan2Class2";
}
}
```

Now, the DHCP server will allocate addresses for the network nodes from Switch1 which are relayed by Switch3 within the range of 192.168.102.21 ~ 192.168.102.50, and allocate addresses for the network nodes from Switch1 within the range of 192.168.102.51 ~ 192.168.102.80.

25.4 DHCP option 82 Troubleshooting

DHCP option 82 is implemented as a sub-function module of DHCP Relay Agent. Before using it, users should make sure that the DHCP Relay Agent is configured correctly.

DHCP option 82 needs the DHCP Relay Agent and the DHCP server cooperate to finish the task of allocating IP addresses. The DHCP server should set allocating policy correctly depending on the network topology of the DHCP Relay Agent, or, even the Relay Agent can

operate normally, the allocation of addresses will fail. When there is more than one kind of Relay Agent, please pay attention to the retransmitting policy of the interface DHCP request messages.

To implement the option 82 function of DHCP Relay Agent, the “debug dhcp relay packet” command can be used during the operating procedure, including adding the contents of option 82, the retransmitting policy adopted, the option 82 contents of the server peeled by the Relay Agent and etc., such information can help users to do troubleshooting.

To implement the option 82 function of DHCP server, the “debug ip dhcp server packet” command can be used during the operating procedure to display the procedure of data packets processing of the server, including displaying the identified option 82 information of the request message and the option 82 information returned by the reply message.

Chapter 26 DHCPv6 option37, 38

26.1 Introduction to DHCPv6 option37, 38

DHCPv6 (Dynamic Host Configuration Protocol for IPv6) is designed for IPv6 address scheme and is used for assigning IPv6 prefixes, IPv6 addresses and other configuration parameters to hosts.

When DHCPv6 client wants to request address and configure parameter of DHCPv6 server from different link, it needs to communicate with server through DHCPv6 relay agent. DHCPv6 message received by relay agent node is reencapsulated to be relay-forward packets and they are forwarded to the server which sends the relay-reply packets to DHCPv6 relay agent node in different link, after that, relay agent node restores DHCPv6 message to DHCPv6 client to finish communication between client and server.

There are some problems when using DHCPv6 relay agent, for example: How to assign IP address in the fixed range to the specifiec users? How to avoid illegal DHCPv6 client to forge IP address exhaust attack triggered by MAC address fields of DHCPv6 packets? How to avoid illegal DHCPv6 client to trigger deny service attack through using MAC address of other legal clients? Therefore, IETF set rfc4649 and rfc4580, i.e. DHCPv6 option 37 and option 38 to solve these problems.

DHCPv6 option 37 and option 38 is similar to DHCP option 82. When DHCPv6 client sends request packets to DHCPv6 server though DHCPv6 relay agent, if DHCPv6 relay agent supports option 37 and option 38, they will be added to request packets. For the respond packets of server, option 37 and option 38 are meaningless and are peeled from the respond packets. Therefore, the application of option 37 and option 38 is transparent for client.

DHCPv6 server can authenticate identity of DHCPv6 client and DHCPv6 relay device by option 37 and option 38, assign and manage client address neatly through configuring the assign policy, prevent DHCPv6 attack availably according to the inclusive client information, such as forging MAC address fields of DHCPv6 packets to trigger IP address exhaust attack. Since server can identify multiple request packets from the same access port, it can assign the address number through policy limit to avoid address exhaust. However, rfc4649 and rfc4580 do not set how to use opton 37 and option 38 for DHCPv6 server, users can use it neatly according to their own demand.

26.2 DHCPv6 option37, 38 Configuration Task List

1. Dhcpv6 snooping option basic functions configuration
2. Dhcpv6 relay option basic functions configuration
3. Dhcpv6 server option basic functions configuration

1.DHCPv6 snooping option basic functions configuration

Command	Description
Global mode	
ipv6 dhcp snooping remote-id option **no ipv6 dhcp snooping remote-id option**	This command enables DHCPv6 SNOOPING to support option 37 option, no command disables it.
ipv6 dhcp snooping subscriber-id option **no ipv6 dhcp snooping subscriber-id option**	This command enables DHCPv6 SNOOPING to support option 38 option, no command disables it.
ipv6 dhcp snooping remote-id policy {drop | keep | replace} **no ipv6 dhcp snooping remote-id policy**	This command is used to configure the reforward policy of the system when receiving DHCPv6 packets with option 37, which can be: **drop,** the system simply discards it with option 37; **keep,** the system keeps option 37 unchanged and forwards the packet to the server; **replace,** the system replaces option 37 of current packet with its own before forwarding it to the server. no command configures the reforward policy of DHCPv6 packets with option 37 as replace.
ipv6 dhcp snooping subscriber-id policy {drop | keep | replace} **no ipv6 dhcp snooping subscriber-id policy**	This command is used to configure the reforward policy of the system when receiving DHCPv6 packets with option 38, which can be: **drop,** the system simply discards it with option 38; **keep,** the system keeps option 38 unchanged and forwards the packet to the server; **replace,** the system replaces option 38 of current packet with its own before forwarding it to the server. no command configures the reforward policy of DHCPv6 packets with option 38 as replace.
ipv6 dhcp snooping subscriber-id select (sp | sv | pv | spv) delimiter WORD (delimiter WORD |) **no ipv6 dhcp snooping subscriber-id**	Configures user configuration options to generate subscriber-id, no command restores to its original default configuration, i.e. enterprise number together with vlan MAC.

select delimiter	
ipv6 dhcp snooping subscriber-id select (sp|sv|pv|spv) delimiter WORD (delimiter WORD |) **no ipv6 dhcp snooping subscriber-id select delimiter**	Configures user configuration options to generate subscriber-id. The no command restores to its original default configuration, i.e. vlan name together with port name.
Port mode	
ipv6 dhcp snooping remote-id <remote-id> **no ipv6 dhcp snooping remote-id**	This command is used to set the form of adding option 37 in received DHCPv6 request packets, of which <remote-id> is the content of remote-id in user-defined option 37 and it is a string with a length of less than 128. The no operation restores remote-id in option 37 to enterprise-number together with vlan MAC address.
ipv6 dhcp snooping subscriber-id <subscriber-id> **no ipv6 dhcp snooping subscriber-id**	This command is used to set the form of adding option 38 in received DHCPv6 request packets, of which <subscriber-id> is the content of subscriber-id in user-defined option 38 and it is a string with a length of less than 128. The no operation restores subscriber-id in option 38 to vlan name together with port name such as "Vlan2+Ethernet1/0/2".

2. DHCPv6 relay option basic functions configuration

Command	Description
Global mode	
ipv6 dhcp relay remote-id option **no ipv6 dhcp relay remote-id option**	This command enables the switch relay to support option 37 and the no form of this command disables it.
ipv6 dhcp relay subscriber-id option **no ipv6 dhcp relay subscriber-id option**	This command enables the switch relay to support the option 38, the no form of this command disables it.
ipv6 dhcp relay remote-id delimiter WORD **no ipv6 dhcp relay remote-id delimiter**	Configures user configuration options to generate remote-id. The no command restores to its original default configuration, i.e. enterprise number together with vlan MAC.

ipv6 dhcp relay subscriber-id select (sp | sv | pv | spv) delimiter WORD (delimiter WORD |) **no ipv6 dhcp relay subscriber-id select delimiter**	Configures user configuration options to generate subscriber-id. The no command restores to its original default configuration, i.e. vlan name together with port name.
Layer3 Interface configuration mode	
ipv6 dhcp relay remote-id <remote-id> **no ipv6 dhcp relay remote-id**	This command is used to set the form of adding option 37 in received DHCPv6 request packets, of which <remote-id> is the content of remote-id in user-defined option 37 and it is a string with a length of less than 128. The no operation restores remote-id in option 37 to enterprise-number together with vlan MAC address.
ipv6 dhcp relay subscriber-id <subscriber-id> **no ipv6 dhcp relay subscriber-id**	This command is used to set the form of adding option 38 in received DHCPv6 request packets, of which <subscriber-id> is the content of subscriber-id in user-defined option 38 and it is a string with a length of less than 128. The no operation restores subscriber-id in option 38 to vlan name together with port name such as "Vlan2+Ethernet1/0".

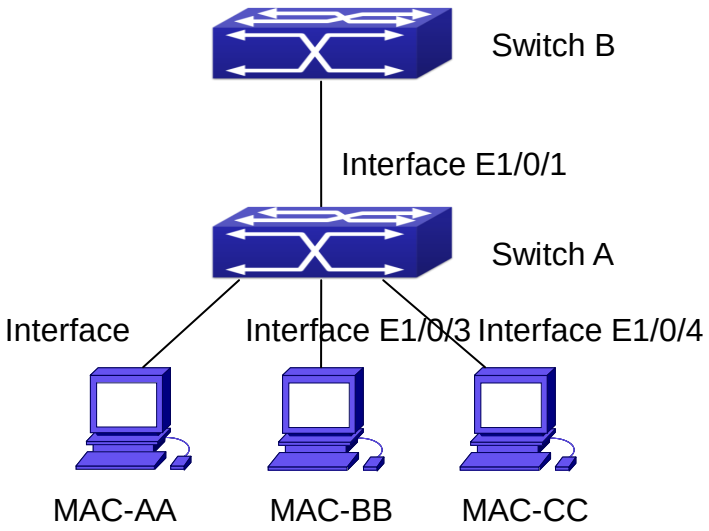
3. Dhcpv6 server option basic functions configuration

Command	Description
Global mode	
ipv6 dhcp server remote-id option **no ipv6 dhcp server remote-id option**	This command enables DHCPv6 server to support the identification of option 37, the no form of this command disables it.
ipv6 dhcp server subscriber-id option **no ipv6 dhcp server subscriber-id option**	This command enables DHCPv6 server to support the identification of option 38, the no form of this command disables it.
ipv6 dhcp use class **no ipv6 dhcp use class**	This command enables DHCPv6 server to support the using of DHCPv6 class during address assignment, the no form of this command disables it without removing the relative DHCPv6 class information that has been configured.

ipv6 dhcp class <class-name> **no ipv6 dhcp class <class-name>**	This command defines a DHCPv6 class and enters DHCPv6 class mode, the no form of this command removes this DHCPv6 class.
Interface configuration mode	
ipv6 dhcp server select relay-forw **no ipv6 dhcp server select relay-forw**	This command enables the DHCPv6 server to support selections when multiple option 37 or option 38 options exist and the option 37 and option 38 of relay-forw in the innermost layer are selected. The no operation of it restores the default configuration, i.e. selecting option 37 and option 38 of the original packets.
IPv6 DHCP Class configuration mode	
{remote-id [*] <remote-id> [*] | subscriber-id [*] <subscriber-id> [*]} **no {remote-id [*] <remote-id> [*] | subscriber-id [*] <subscriber-id> [*]}**	This command configures option 37 and option 38 that match the class in ipv6 dhcp class configuration mode.
DHCPv6 address pool configuration mode	
class <class-name> **no class <class-name>**	This command associates class to address pool in DHCPv6 address pool configuration mode and enters class configuration mode in address pool. Use no command to remove the link.
address range <start-ip> <end-ip> **no address range <start-ip> <end-ip>**	This command is used to set address range for a DHCPv6 class in DHCPv6 address pool configuration mode, the no command is used to remove the addreass range. The prefix/plen form is not supported.

26.3 DHCPv6 option37, 38 Examples

26.3.1 DHCPv6 Snooping option37, 38 Example



DHCPv6 Snooping option schematic

As is shown in the figure above, Mac-AA, Mac-BB and Mac-CC are normal users, connected to untrusted interface 1/2, 1/3 and 1/4 respectively, and they get IP 2010:2, 2010:3 and 2010:4 through DHCPv6 Client; DHCPv6 Server is connected to the trusted interface 1/1. Configure three address assignment policies (CLASS), of which CLASS1 matches option 38, CLASS2 matches option 37 and CLASS3 matches option 37 and option 38. In the address pool EastDormPool, the requests matched with CLASS1, CLASS2 and CLASS3 will be assigned an address ranging from 2001:da8:100:1::2 to 2001:da8:100:1::30, from 2001:da8:100:1::31 to 2001:da8:100:1::60 and from 2001:da8:100:1::61 to2001:da8:100:1::100 respectively; DHCPv6 snooping function is enabled and option 37 and option 38 are configured in Switch A.
Switch A configuration:

```
SwitchA(config)#ipv6 dhcp snooping remote-id option
SwitchA(config)#ipv6 dhcp snooping subscriber-id option
SwitchA(config)#int e 1/0/1
SwitchA(config-if-ethernet1/0/1)#ipv6 dhcp snooping trust
SwitchA(config-if-ethernet1/0/1)#exit
SwitchA(config)#interface vlan 1

SwitchA(config-if-vlan1)#ipv6 address 2001:da8:100:1::1
SwitchA(config-if-vlan1)#exit
SwitchA(config)#interface ethernet 1/0/1-4
SwitchA(config-if-port-range)#switchport access vlan 1
SwitchA(config-if-port-range)#exit
SwitchA(config)#
```

Switch B configuration:

```
SwitchB(config)#service dhcpv6
SwitchB(config)#ipv6 dhcp server remote-id option
SwitchB(config)#ipv6 dhcp server subscriber-id option
SwitchB(config)#ipv6 dhcp pool EastDormPool
SwitchB(dhcpv6-eastdormpool-config)#network-address 2001:da8:100:1::2
2001:da8:100:1::1000
SwitchB(dhcpv6-eastdormpool-config)#dns-server 2001::1
SwitchB(dhcpv6-eastdormpool-config)#domain-name dhcpv6.com
SwitchB(dhcpv6-eastdormpool-config)# excluded-address 2001:da8:100:1::2
SwitchB(dhcpv6-eastdormpool-config)#exit
SwitchB(config)#
SwitchB(config)#ipv6 dhcp class CLASS1
SwitchB(dhcpv6-class-class1-config)#remote-id 00-1f-ce-00-00-01 subscriber-id
vlan1+Ethernet1/0/1
SwitchB(dhcpv6-class-class1-config)#exit
SwitchB(config)#ipv6 dhcp class CLASS2
SwitchB(dhcpv6-class-class2-config)#remote-id 00-1f-ce-00-00-01 subscriber-id
vlan1+Ethernet1/0/2
SwitchB(dhcpv6-class-class2-config)#exit
SwitchB(config)#ipv6 dhcp class CLASS3
SwitchB(dhcpv6-class-class3-config)#remote-id 00-1f-ce-00-00-01 subscriber-id
vlan1+Ethernet1/0/3
SwitchB(dhcpv6-class-class3-config)#exit
SwitchB(config)#ipv6 dhcp pool EastDormPool
SwitchB(dhcpv6-eastdormpool-config)#class CLASS1
SwitchB(dhcpv6-pool-eastdormpool-class-class1-config)#address range 2001:da8:100:1::3
2001:da8:100:1::30
SwitchB(dhcpv6-pool-eastdormpool-class-class1-config)#exit
witchB(dhcpv6-eastdormpool-config)#class CLASS2
SwitchB(dhcpv6-pool-eastdormpool-class-class2-config)#address range 2001:da8:100:1::31
2001:da8:100:1::60
SwitchB(dhcpv6-eastdormpool-config)#class CLASS3
SwitchB(dhcpv6-pool-eastdormpool-class-class3-config)#address range 2001:da8:100:1::61
2001:da8:100:1::100
SwitchB(dhcpv6-pool-eastdormpool-class-class3-config)#exit
SwitchB(dhcpv6-eastdormpool-config)#exit
SwitchB(config)#interface vlan 1
```

SwitchB(config-if-vlan1)#ipv6 address 2001:da8:100:1::2/64
SwitchB(config-if-vlan1)#ipv6 dhcp server EastDormPool
SwitchB(config-if-vlan1)#exit
SwitchB(config)#

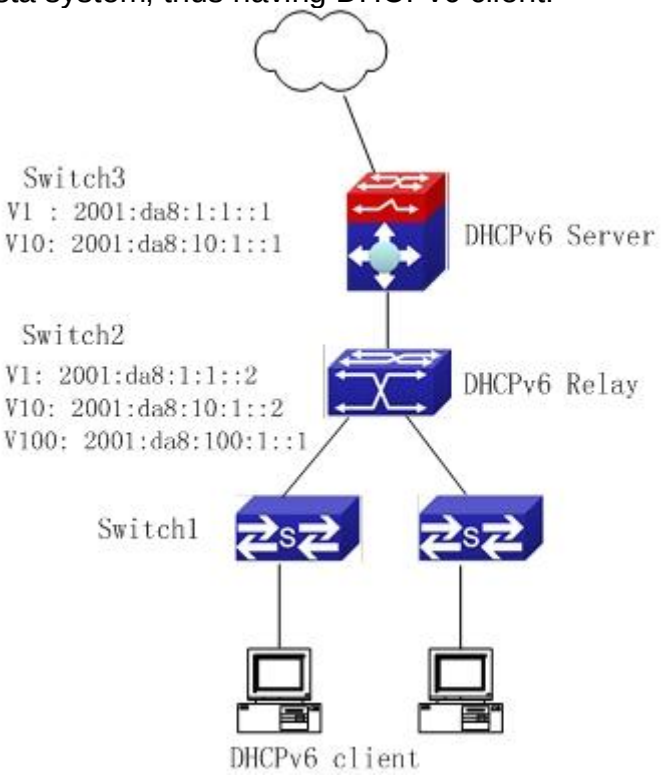
26.3.2 DHCPv6 Relay option37, 38 Example

Example 1:

When deploying IPv6 campus network, DHCPv6 server function of routing device can be used for IPv6 address allocation if special server is used for uniform allocation and management for IPv6 address. DHCPv6 server supports both stateful and stateless DHCPv6.

Network topology:

In access layer, layer2 access device Switch1 connects users in dormitory; in first-level aggregation layer, aggregation device Switch2 is used as DHCPv6 relay agent; in second-level aggregation layer, aggregation device Switch3 is used as DHCPv6 server and connects with backbone network or devices in higher aggregation layer; in user side, PCs are generally loaded with Windows Vista system, thus having DHCPv6 client.



DHCPv6 relay option schematic

Switch2 configuration:

S2(config)#service dhcpv6
S2(config)#ipv6 dhcp relay remote-id option
S2(config)#ipv6 dhcp relay subscriber-id option
S2(config)#vlan 10

```
S2(config-vlan10)#int vlan 10
S2(config-if-vlan10)#ipv6 address 2001:da8:1:::2/64
S2(config-if-vlan10)#ipv6 dhcp relay destination 2001:da8:10:1::1
S2(config-if-vlan10)#exit
S2(config)#
```

26.4 DHCPv6 option37, 38 Troubleshooting

Request packets sent by DHCPv6 client are multicast packets received by the device within its VLAN, if DHCPv6 server wants to receive the packets from client, DHCPv6 client and DHCPv6 server must be in the same VLAN, otherwise it needs to use DHCPv6 relay.

Snooping option37,38 can process one of the following operations for DHCPv6 request packets with option37,38: replace the original option37,38 with its own; discard the packets with option37,38; do not execute adding, discarding or forwarding operation. Therefore, please check policy configuration of snooping option37,38 on second device when obtaining the false address or no address is obtained according to option37,38.

DHCPv6 server obtains option37,38 of the packets from client by default, if no, it will obtain option37,38 of the packet sent by relay.

DHCPv6 server only checks whether the first DHCPv6 relay adds option37,38 that means only option37,38 of the innermost relay-forw is valid in relay packets.

Chapter 27 DHCP Snooping Configuration

27.1 Introduction to DHCP Snooping

DHCP Snooping means that the switch monitors the IP-getting process of DHCP CLIENT via DHCP protocol. It prevents DHCP attacks and illegal DHCP SERVER by setting trust ports and untrust ports. And the DHCP messages from trust ports can be forwarded without being verified. In typical settings, trust ports are used to connect DHCP SERVER or DHCP RELAY Proxy, and untrust ports are used to connect DHCP CLINET. The switch will forward the DCHP request messages from untrust ports, but not DHCP reply ones. If any DHCP reply messages is received from a untrust port, besides giving an alarm, the switch will also implement designated actions on the port according to settings, such as “shutdown”, or distributing a “blackhole”. If DHCP Snooping binding is enabled, the switch will save binding information (including its MAC address, IP address, IP lease, VLAN number and port number) of each DHCP CLINET on untrust ports in DHCP snooping binding table With such information, DHCP Snooping can combine modules like dot1x and ARP, or implement user-access-control independently.

Defense against Fake DHCP Server: once the switch intercepts the DHCP Server reply packets(including DHCPOFFER, DHCPACK, and DHCPNAK) , it will alarm and respond according to the situation(shutdown the port or send Black hole) .
Defense against DHCP over load attacks: To avoid too many DHCP messages attacking CPU, users should limit the DHCP speed of receiving packets on trusted and non-trusted ports.
Record the binding data of DHCP: DHCP SNOOPING will record the binding data allocated by DHCP SERVER while forwarding DHCP messages, it can also upload the binding data to the specified server to backup it. The binding data is mainly used to configure the dynamic users of dot1x user based ports. Please refer to the chapter called“dot1x configuration” to find more about the usage of dot1x use-based mode.
Add binding ARP: DHCP SNOOPING can add static binding ARP according to the binding data after capturing binding data, thus to avoid ARP cheating.
Add trusted users: DHCP SNOOPING can add trusted user list entries according to the parameters in binding data after capturing binding data; thus these users can access all resources without DOT1X authentication.
Automatic Recovery: A while after the switch shut down the port or send blockhole, it should automatically recover the communication of the port or source MAC and send information to Log Server via syslog.
LOG Function: When the switch discovers abnormal received packets or automatically recovers, it should send syslog information to Log Server.

The Encryption of Private Messages: The communication between the switch and the inner network security management system TrustView uses private messages. And the users can encrypt those messages of version 2.

Add authentication option82 Function: It is used with dot1x dhcpoption82 authentication mode. Different option 82 will be added in DHCP messages according to user's authentication status.

27.2 DHCP Snooping Configuration Task Sequence

1. Enable DHCP Snooping
2. Enable DHCP Snooping binding function
3. Enable DHCP Snooping binding ARP function
4. Enable DHCP Snooping option82 function
5. Set the private packet version
6. Set DES encrypted key for private packets
7. Set helper server address
8. Set trusted ports
9. Enable DHCP Snooping binding DOT1X function
10. Enable DHCP Snooping binding USER function
11. Adding static list entries function
12. Set defense actions
13. Set rate limitation of DHCP messages
14. Enable the debug switch
15. Configure DHCP Snooping option 82 attributes

1. Enable DHCP Snooping

Command	Explanation
Globe mode	
ip dhcp snooping enable **no ip dhcp snooping enable**	Enable or disable the DHCP snooping function.

2. Enable DHCP Snooping binding

Command	Explanation
Globe mode	
ip dhcp snooping binding enable **no ip dhcp snooping binding enable**	Enable or disable the DHCP snooping binding function.

3. Enable DHCP Snooping binding ARP function

Command	Explanation
Globe mode	
ip dhcp snooping binding arp **no ip dhcp snooping binding arp**	Enable or disable the dhcp snooping binding ARP function.

4. Enable DHCP Snooping option82 function

Command	Explanation
Globe mode	
ip dhcp snooping information enable **no ip dhcp snooping information enable**	Enable/disable DHCP Snooping option 82 function.

5. Set the private packet version

Command	Explanation
Globe mode	
ip user private packet version two **no ip user private packet version two**	To configure/delete the private packet version.

6. Set DES encrypted key for private packets

Command	Explanation
Globe mode	
enable trustview key 0/7 <*password*> **no enable trustview key**	To configure/delete DES encrypted key for private packets.

7. Set helper server address

Command	Explanation
Globe mode	
ip user helper-address A.B.C.D [port <udpport>] source <ipAddr> (secondary|) **no ip user helper-address (secondary|)**	Set or delete helper server address.

8. Set trusted ports

Command	Explanation
Port mode	
ip dhcp snooping trust	Set or delete the DHCP snooping trust attributes of

no ip dhcp snooping trust	ports.

9. Enable DHCP SNOOPING binding DOT1X function

Command	Explanation
Port mode	
ip dhcp snooping binding dot1x **no ip dhcp snooping binding dot1x**	Enable or disable the DHCP snooping binding dot1x function.

10. Enable or disable the DHCP SNOOPING binding USER function

Command	Explanation
Port mode	
ip dhcp snooping binding user-control **no ip dhcp snooping binding user-control**	Enable or disable the DHCP snooping binding user function.

11. Add static binding information

Command	Explanation
Globe mode	
ip dhcp snooping binding user *<mac>* address *<ipAddr>* *<mask>* vlan *<vid>* interface (ethernet|) *<ifname>* **no ip dhcp snooping binding user *<mac>* interface (ethernet|) *<ifname>***	Add/delete DHCP snooping static binding list entries.

12. Set defense actions

Command	Explanation
Port mode	
ip dhcp snooping action {shutdown|blackhole} [recovery *<second>*] **no ip dhcp snooping action**	Set or delete the DHCP snooping automatic defense actions of ports.

13. Set rate limitation of data transmission

Command	Explanation
Globe mode	

ip dhcp snooping limit-rate <*pps*> **no ip dhcp snooping limit-rate**	Set rate limitation of the transmission of DHCP snooping messages.

14. Enable the debug switch

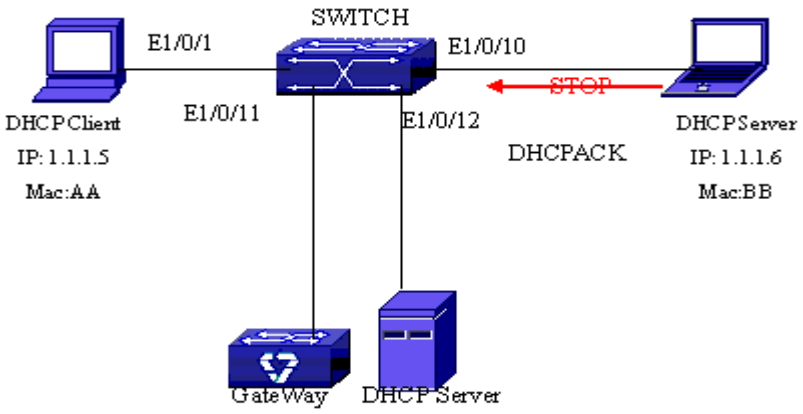
Command	Explanation
Admin mode	
debug ip dhcp snooping packet **debug ip dhcp snooping event** **debug ip dhcp snooping update** **debug ip dhcp snooping binding**	Please refer to the chapter on system troubleshooting.

15. Configure DHCP Snooping option 82 attributes

Command	Explanation
Globe mode	
ip dhcp snooping information option subscriber-id format {hex | acsii | vs-hp}	This command is used to set subscriber-id format of DHCP snooping option82.
ip dhcp snooping information option remote-id {standard | <*remote-id*>} **no ip dhcp snooping information option remote-id**	Set the suboption2 (remote ID option) content of option 82 added by DHCP request packets (they are received by the port). The no command sets the additive suboption2 (remote ID option) format of option 82 as standard.
ip dhcp snooping information option allow-untrusted **no ip dhcp snooping information option allow-untrusted**	This command is used to set that allow untrusted ports of DHCP snooping to receive DHCP packets with option82 option. When disabling this command, all untrusted ports will drop DHCP packets with option82 option.
ip dhcp snooping information option delimiter [colon | dot | slash | space] **no ip dhcp snooping information option delimiter**	Set the delimiter of each parameter for suboption of option82 in global mode, no command restores the delimiter as slash.
ip dhcp snooping information option self-defined remote-id {hostname | mac | string WORD} **no ip dhcp snooping information option self-defined remote-id**	Set creation method for option82, users can define the parameters of remote-id suboption by themselves.
ip dhcp snooping information option self-defined remote-id format [ascii | hex]	Set self-defined format of remote-id for snooping option82.

ip dhcp snooping information option self-defined subscriber-id {vlan | port | id (switch-id (mac | hostname)| remote-mac) | string WORD} **no ip dhcp snooping information option type self-defined subscriber-id**	Set creation method for option82, users can define the parameters of circute-id suboption by themselves.
ip dhcp snooping information option self-defined subscriber-id format [ascii | hex]	Set self-defined format of circuit-id for snooping option82.
Port mode	
ip dhcp snooping information option subscriber-id {standard | <*circuit-id*>} **no ip dhcp snooping information option subscriber-id**	Set the suboption1 (circuit ID option) content of option 82 added by DHCP request packets (they are received by the port). The no command sets the additive suboption1 (circuit ID option) format of option 82 as standard.

27.3 DHCP Snooping Typical Application



Typical usage

As showed in the above chart, Mac-AA device is the normal user, connected to the non-trusted port 1/0/1 of the switch. It operates via DHCP Client, IP 1.1.1.5; DHCP Server and GateWay are connected to the trusted ports 1/0/11 and 1/0/12 of the switch; the malicious user Mac-BB is connected to the non-trusted port 1/0/10, trying to fake a DHCP Server(by sending DHCPACK) . Setting DHCP Snooping on the switch will effectively detect and block this kind of network attack.

Configuration sequence is:

switch#

switch#config

```
switch(config)#ip dhcp snooping enable
switch(config)#interface ethernet 1/0/11
switch(Config-If-Ethernet1/0/11)#ip dhcp snooping trust
switch(Config-If-Ethernet1/0/11)#exit
switch(config)#interface ethernet 1/0/12
switch(Config-If-Ethernet1/0/12)#ip dhcp snooping trust
switch(Config-If-Ethernet1/0/12)#exit
switch(config)#interface ethernet 1/0/1-10
switch(Config-Port-Range)#ip dhcp snooping action shutdown
switch(Config-Port-Range)#
```

27.4 DHCP Snooping Troubleshooting Help

27.4.1 Monitor and Debug Information

The “debug ip dhcp snooping” command can be used to monitor the debug information.

27.4.2 DHCP Snooping Troubleshooting Help

If there is any problem happens when using DHCP Snooping function, please check if the problem is caused by the following reasons:
Check that whether the global DHCP Snooping is enabled;
If the port does not react to invalid DHCP Server packets, please check that whether the port is set as a non-trusted port of DHCP Snooping.

Chapter 28 IPv4 Multicast Protocol

28.1 IPv4 Multicast Protocol Overview

This chapter will give an introduction to the configuration of IPv4 Multicast Protocol. All IPs in this chapter are IPv4.

28.1.1 Introduction to Multicast

Various transmission modes can be adopted when the destination of packet (including data, sound and video) transmission is the minority users in the network. One way is to use Unicast mode, i.e. to set up a separate data transmission path for each user; or, to use Broadcast mode, which is to send messages to all users in the network, and they will receive the Broadcast messages no matter they need or not. For example, if there are 200 users in a network who want to receive the same packet, then the traditional solution is to send this packet for 200 times separately via Unicast to guarantee the users who need the data can get all data wanted, or send the data in the entire domain via Broadcast. Transferring the data in the whole range of network .The users who need these data can get directly from the network. Both modes waste a great deal of valuable bandwidth resource, and furthermore, Broadcast mode goes against the security and secrecy.

The emergence of IP Multicast technology solved this problem in time. The Multicast source only sends out the message once, Multicast Routing Protocol sets up tree-routing for Multicast data packet, and then the transferred packet just starts to be duplicated and distributed in the bifurcate crossing as far as possible. Thus the packet can be sent to every user who needs it accurately and effectively.

It should be noticed that it is not necessary for Multicast source to join in Multicast group. It sends data to some Multicast groups, but it is not necessarily a receiver of the group itself. There can be more than one source sending packets to a Multicast group simultaneously. There may exist routers in the network which do not support Multicast, but a Multicast router can encapsulate the Multicast packets into Unicast IP packets with tunnel mode to send them to the Multicast router next to it, which will take off the Unicast IP header and continue the Multicast transmission process, thus a big alteration of network structure is avoided. The primary advantages of Multicast are:

Enhance efficiency: reduce network traffic, lighten the load of server and CPU

Optimize performance: reduce redundant traffic

Distributed application: Enable Multipoint Application

28.1.2 Multicast Address

The destination address of Multicast message uses class D IP address with range from 224.0.0.0 to 239.255.255.255. D class address can not appear in the source IP address field of an IP message. In the process of Unicast data transmission, the transmission path of a data packet is from source address routing to destination address, and the transmission is performed with hop-by-hop principle. However, in IP Multicast environment, the destination addresses is a group instead of a single one, they form a group address. All message receivers will join in a group, and once they do, the data flowing to the group address will be sent to the receivers immediately and all members in the group will receive the data packets. The members in a Multicast group are dynamic, the hosts can join and leave the Multicast group at any time.

Multicast group can be permanent or temporary. Some of the Multicast group addresses are assigned officially; they are called Permanent Multicast Group. Permanent Multicast Group keeps its IP address fixed but its member structure can vary within. The member amount of Permanent Multicast Group can be arbitrary, even zero. The IP Multicast addresses which are not kept for use by Permanent Multicast Group can be utilized by temporary Multicast groups.

224.0.0.0~224.0.0.255 are reserved Multicast addresses (Permanent Group Address), address 224.0.0.0 is reserved but not assigned, and other addresses are used by Routing Protocol; 224.0.1.0~238.255.255.255 are Multicast addresses available to users(Temporary Group Address) and are valid in the entire domain of the network; 239.0.0.0~239.255.255.255 are local management Multicast addresses, which are valid only in specific local domain. Frequently used reserved multicast address list is as follows:

Benchmark address (reserved)
224.0.0.1 Address of all hosts
224.0.0.2 Address of all Multicast Routers
224.0.0.3 Unassigned
224.0.0.4 DVMRP Router
224.0.0.5 OSPF Router
224.0.0.6 OSPF DR
224.0.0.7 ST Router
224.0.0.8 ST host
224.0.0.9 RIP-2 Router
224.0.0.10 IGRP Router
224.0.0.11 Active Agent
224.0.0.12 DHCP Server/Relay Agent
224.0.0.13 All PIM Routers
224.0.0.14 RSVP Encapsulation
224.0.0.15 All CBT Routers
224.0.0.16 Specified SBM

224.0.0.17 All SBMS
224.0.0.18 VRRP
224.0.0.22 IGMP
When Ethernet transmits Unicast IP messages, the destination MAC address it uses is the receiver's MAC address. But in transmitting Multicast packets, the transmission destination is not a specific receiver any more, but a group with uncertain members, thus Multicast MAC address is used. Multicast MAC address is corresponding to Multicast IP address. It is prescribed in IANA (Internet Assigned Number Authority) that the higher 25 bits in Multicast MAC address is 0x01005e, and the lower 23bits in MAC address is the lower 23bits in Multicast IP address.
Since only 23bits out of the lower 28bits in IP Multicast address are mapped into MAC address, therefore there are 32 IP Multicast addresses which are mapped into the same MAC address.

28.1.3 IP Multicast Packet Transmission

In Multicast mode, the source host sends packets to the host group indicated by the Multicast group address in the destination address field of IP data packet. Unlike Unicast mode, Multicast data packet must be forwarded to a number of external interfaces to be sent to all receiver sites in Multicast mode, thus Multicast transmission procedure is more complicated than Unicast transmission procedure.
In order to guarantee that all Multicast packets get to the router via the shortest path, the receipt interface of the Multicast packet must be checked in some certain way based on Unicast router table; this checking mechanism is the basis for most Multicast Routing Protocol to forward in Multicast mode --- RPF (Reverse Path Forwarding) check. Multicast router makes use of the impressed packet source address to query Unicast Router Table or independent Multicast Router Table to determine if the packet ingress interface is on the shortest path from receipt site to source address. If shortest path Tree is used, then the source address is the address of source host which sends Multicast Data Packets; if Shared Tree is used, then the source address is the address of the root of the Shared-Tree. When Multicast data packet gets to the router, if RPF check passes, then the data packet is forwarded according to Multicast forward item, and the data packet will be discarded else wise.

28.1.4 IP Multicast Application

IP Multicast technology has effectively solved the problem of sending in single point and receiving in multipoint. It has achieved the effective data transmission from a point to multiple points, saved a great deal of network bandwidth and reduced network load. Making use of the Multicast property of network, some new value-added operations can be supplied conveniently. In Information Service areas such as online living broadcast, network TV, remote education, remote medicine, real time video/audio meeting, the following applications may be supplied:
Application of Multimedia and Streaming Media

Data repository, finance application (stock) etc
Any data distribution application of “one point to multiple points”
In the situation of more and more multimedia operations in IP network, Multicast has tremendous market potential and Multicast operation will be generalized and popularized.

28.2 PIM-DM

28.2.1 Introduction to PIM-DM

PIM-DM(Protocol Independent Multicast, Dense Mode) is a Multicast Routing Protocol in dense mode which applies to small network. The members of multicast group are relatively dense under this kind of network environment.
The working process of PIM-DM can be summarized as: Neighbor Discovery, Flooding & Prune, and Graft.
1. Neigh hour Discovery
After PIM-DM router is enabled, Hello message is required to discover neighbors. The network nodes which run PIM-DM use Hello message to contact each other. PIM-DM Hello message is sent periodically.
2. Flooding & Prune of process
PIM-DM assumes all hosts on the network are ready to receive Multicast data. When some Multicast Source begins to send data to a Multicast Group G, after receiving the Multicast packet, the router will make RPF check first according to the Unicast table. If the check passes, the router will create a (S, G) table entry and transmit the Multicast packet to all downstream PIM-DM nodes on the network (Flooding). If the RPF check fails, i.e. the Multicast packet is input from the incorrect interface, and then the message is discarded. After this procedure, in the PIM-DM Multicast domain, every node will create a (S, G) table entry. If there is no Multicast group member in the downstream nodes, then a Prune message is sent to upstream nodes to notify them not to transmit data of this Multicast group any more. After receiving Prune message, the upstream nodes will delete the corresponding interface from the output interface list to which their Multicast transmission table entry (S, G) corresponds. Thus a SPT(Shortest Path Tree, SPT) tree with source S as root is created. The Prune process is initiated by leaf router first.
The process above is called Flooding & Prune process. Each pruned node also provides time-out mechanics at the same time. When Prune is timed-out, the router will restart Flooding & Prune process. The PIM-DM Flooding & Prune is periodically processed.
3. RPF Check
With RPF Check, PIM-DM makes use of existing Unicast routing table to establish a Multicast transmission tree initiating from data source. When a Multicast packet arrives, the router will determine whether the coming path is correct first. If the arrival interface is the interface

connected to Multicast source indicated by Unicast routing, then this Multicast packet is considered to be from the correct path. Otherwise the Multicast packet is to be discarded as redundant message. The Unicast routing message used as path judgment can root in any Unicast Routing Protocol, such as messages found by RIP, OSPF, etc. It doesn't rely on any specific Unicast Routing Protocol.

4. Assert Mechanism

If each of two Multicast routers A and B on the same LAN segment has a receiving route respectively and both will transmit the Multicast packet to the LAN after receiving the Multicast data packet sent by the Multicast Source S, then the downstream node Multicast router C will receive two exactly same Multicast packets. The router needs to choose a unique transmitter through Assert mechanism after it detects this situation. An optimal transmission path is selected through sending out Assert packet. If the priority and cost of two or more path are same, then the node with larger IP address is taken as the upstream neighbor of the (S, G) entry and in charge of the transmission of the (S, G) Multicast packet.

5. Graft

When the pruned downstream node needs to recover to transmission status, this node uses Graft Packet to notify upstream nodes to restore multicast data transmission.

28.2.2 PIM-DM Configuration Task List

Enable PIM-DM (Required)
Configure static multicast routing entries(Optional)
Configure additional PIM-DM parameters(Optional)
Configure the interval for PIM-DM hello messages
Configure the interval for state-refresh messages
Configure the boundary interfaces
Configure the management boundary
Disable PIM-DM protocol

1. Enable the PIM-DM protocol

When configuring the PIM-DM protocol on QTECH series Layer 3 switches, PIM multicasting should be enabled globally, then PIM-DM can be enabled for specific interfaces.

Command	Explanation
Global Mode	
ip pim multicast-routing **no ip pim multicast-routing**	To enable PIM-DM globally for all the interfaces (However, in order to make PIM-DM work for specific interfaces, the following command should be issued).

And then turn on PIM-SM switch on the interface

Command	Explanation
Interface Configuration Mode	

ip pim dense-mode	To enable PIM-DM protocol for the specified interface.(Required)

2. Configure static multicast routing entries

Command	Explanation
Global Configuration Mode	
ip mroute <A.B.C.D> <A.B.C.D> <ifname> <.ifname> **no ip mroute <*A.B.C.D*> <*A.B.C.D*> [<*ifname*> <.*ifname*>]**	To configure a static multicast routing entry. The no form of this command will remove the specified entry.

3. Configure additional PIM-DM parameters

Configure the interval for PIM-DM hello messages

Command	Explanation
Interface Configuration Mode	
ip pim hello-interval < *interval*> **no ip pim hello-interval**	To configure the interval for PIM-DM hello messages. The no form of this command will restore the interval to the default value.

Configure the interval for state-refresh messages

Command	Explanation
Interface Configuration Mode	
ip pim state-refresh origination-interval **no ip pim state-refresh origination-interval**	To configure the interval for sending PIM-DM state-refresh packets. The no form of this command will restore the default value.

Configure the boundary interfaces

Command	Explanation
Interface Configuration Mode	
ip pim bsr-border **no ip pim bsr-border**	To configure the interface as the boundary of PIM-DM protocol. On the boundary interface, BSR messages will not be sent or received. The network connected the interface is considered as directly connected network. The no form of this command will remove the configuration.

Configure the management boundary

Command	Explanation
Interface Configuration Mode	

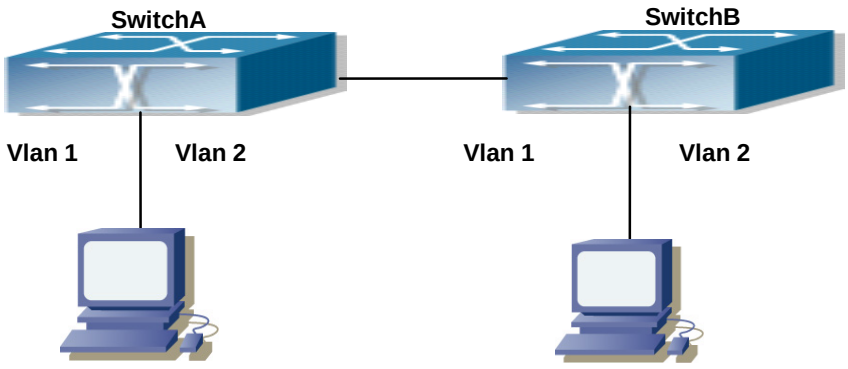
ip pim scope-border <*1-99* >|<*acl_name*> **no ip pim scope-border**	To configure PIM-DM management boundary for the interface and apply ACL for the management boundary. With default settings, 239.0.0.0/8 is considered as the scope of the management group. If ACL is configured, then the scope specified by ACL permit command is the scope of the management group. The no form of this command will remove the configuration.

4. Disable PIM-DM protocol

Command	Explanation
Interface Configuration Mode	
no ip pim dense-mode	To disable the PIM-DM protocol for the interface.
Global Configuration Mode	
no ip pim multicast-routing	To disable PIM-DM globally.

28.2.3 PIM-DM Configuration Examples

As shown in the following figure, add the Ethernet interfaces of Switch A and Switch B to corresponding vlan, and enable PIM-DM Protocol on each vlan interface.



PIM-DM Typical Environment

The configuration procedure for SwitchA and SwitchB is as follows:

(1) Configure SwitchA:

Switch(config)#ip pim multicast-routing

Switch(config)#interface vlan 1

Switch(Config-if-Vlan1)# ip address 10.1.1.1 255.255.255.0

Switch(Config-if-Vlan1)# ip pim dense-mode

Switch(Config-if-Vlan1)#exit

Switch(config)#interface vlan2

Switch(Config-if-Vlan2)# ip address 12.1.1.1 255.255.255.0

Switch(Config-if-Vlan2)# ip pim dense-mode

(2) Configure SwitchB:

Switch(config)#ip pim multicast-routing
Switch(config)#interface vlan 1
Switch(Config-if-Vlan1)# ip address 12.1.1.2 255.255.255.0
Switch(Config-if-Vlan1)# ip pim dense-mode
Switch(Config-if-Vlan1)#exit
Switch(config)#interface vlan 2
Switch(Config-if-Vlan2)# ip address 20.1.1.1 255.255.255.0
Switch(Config-if-Vlan2)# ip pim dense-mode
At the same time, you should pay attention to the configuration of Unicast Routing Protocol, assure that each device can communicate with each other in the network layer, and be able to implement dynamic routing update in virtue of Unicast Routing Protocol.

28.2.4 PIM-DM Troubleshooting

In configuring and using PIM-DM Protocol, PIM-DM Protocol might not operate normally caused by physical connection or incorrect configuration. Therefore, the user should pay attention to the following issues:
To assure that physical connection is correct
To assure the Protocol of Interface and Link is UP (use show interface command)
To assure PIM Protocol is enabled in Global Mode (use ipv6 pim multicast-routing)
Enable PIM-DM Protocol on the interface (use ipv6 pim dense-mode command)
Multicast Protocol requires RPF Check using Unicast routing; therefore the correctness of Unicast routing must be assured beforehand
If all attempts including Check are made but the problems on PIM-DM can't be solved yet, then use debug commands such as debug pim please, and then copy DEBUG information in 3 minutes and send to Technology Service Center.

28.3 PIM-SM

28.3.1 Introduction to PIM-SM

PIM-SM(Protocol Independent Multicast, Sparse Mode) is Protocol Independent Multicast Sparse Mode. It is a Multicast Routing Protocol in Sparse Mode and mainly used in big scale network with group members distributed relatively sparse and wide-spread. Unlike the Flooding & Prune of Dense Mode, PIM-SM Protocol assumes no host needs receiving Multicast data packets. PIM-SM router transmits Multicast Data Packets to a host only if it presents explicit requirement.
By setting RP (Rendezvous Point) and BSR (Bootstrap Router), PIM-SM announce Multicast packet to all PIM-SM routers and establish RPT (RP-rooted shared tree) based on RP using Join/Prune message of routers. Consequently the network bandwidth occupied by data

packets and message control is cut down and the transaction cost of routers decreases. Multicast data get to the network segment where the Multicast group members are located along the shared tree flow. When the data traffic reaches a certain amount, Multicast data stream can be switched to the shortest path tree SPT based on the source to reduce network delay. PIM-SM doesn't rely on any specific Unicast Routing Protocol but make RPF Check using existing Unicast routing table.

1. PIM-SM Working Principle

The central working processes of PIM-SM are: Neighbor Discovery, Generation of RP Shared Tree (RPT), Multicast source registration, SPT Switch, etc. We won't describe the mechanism of Neighbor Discovery here since it is same as that of PIM-DM.

Generation of RP Shared Tree (RPT)

When a host joins a Multicast Group G, the leaf router that is connected to this host directly finds out through IGMP message that there is a receiver of Multicast Group G, then it works out the corresponding Rendezvous Point RP for Multicast Group G, and send join message to upper lever nodes in RP direction. Every router on the way from the leaf router to RP will generate a (*, G) table entry, where a message from any source to Multicast group applies to this entry. When RP receives the message sent to Multicast Group G, the message will get to the leaf router along the set up path and reach the host. In this way the RPT with RP as root is generated.

Multicast Source Registration

When a Multicast Source S sends a Multicast packet to Multicast Group G, the PIM-SM Multicast router connected to it directly will take charge of encapsulating the Multicast packet into registered message and unicast it to corresponding RP. If there are more than one PIM-SM Multicast routers on a network segment, then DR (Designated Router) takes charge of sending the Multicast packet.

SPT Switch

When the Multicast router finds that the rate of the Multicast packet from RP with destination address G exceeds threshold, the Multicast router will send Join message to the next upper lever nodes in the source direction, which results in the switch from RPT to SPT.

2. Preparation before PIM-SM configuration

Configuration Candidate RP

More than one RPs (candidate RP) can exist in PIM-SM network and each C-RP (Candidate RP) takes charge of transmitting Multicast packets with destination address in a certain range. To configure more than one candidate RPs can implement RP load share. No master or slave is differentiated among RPs. All Multicast routers work out the RP corresponding to some Multicast group based on the same algorithm after receiving the candidate RP message announced by BSR.

Note that one RP can serve more than one Multicast groups and all Multicast groups. Each Multicast group can only correspond to one unique RP at any moment. It can't correspond to

more than one RP at the same time.

Configure BSR

BSR is the management center of PIMSM network. It is in charge of collecting messages sent by candidate RPs and broadcast them.

Only one BSR can exist within a network, but more than one C-BSR (Candidate-BSR) can be configured. In this way, if some BSR goes wrong, it can switch to another. C-BSRs elect BSR automatically.

28.3.2 PIM-SM Configuration Task List

Enable PIM-SM (Required)

Configure static multicast routing entries (Optional)

Configure additional parameters for PIM-SM (Optional)

Configure parameters for PIM-SM interfaces

Configure the interval for PIM-SM hello messages

Configure the hold time for PIM-SM hello messages

Configure ACL for PIM-SM neighbors

Configure the interface as the boundary interface of the PIM-SM protocol

Configure the interface as the management boundary of the PIM-SM protocol

Configure global PIM-SM parameters

Configure the switch as a candidate BSR

Configure the switch as a candidate RP

Configure static RP

Configure the cache time of kernel multicast route

Disable PIM-SM Protocol

1. Enable PIM-SM Protocol

The PIM-SM protocol can be enabled on QTECH series Layer 3 switches by enabling PIM in global configuration mode and then enabling PIM-SM for specific interfaces in the interface configuration mode.

Command	Explanation
Global Mode	
ip pim multicast-routing	To enable the PIM-SM protocol for all the interfaces (However, in order to make PIM-SM work for specific interfaces, the following command should be issued).(Required)

And then turn on PIM-SM switch on the interface

Command	Explanation
Interface Configuration Mode	
ip pim sparse-mode	Enable PIM-SM Protocol of the interface.

	(Required).

2. Configure static multicast routing entries

Command	Explanation
Global Configuration Mode	
ip mroute <A.B.C.D> <A.B.C.D> <ifname> <.ifname> **no ip mroute <*A.B.C.D*> <*A.B.C.D*> [<*ifname*> <*.ifname*>]**	To configure a static multicast routing entry. The no form of this command will remove the specified static multicast routing entry.

3. Configure additional parameters for PIM-SM

Configure parameters for PIM-SM interfaces

Configure the interval for PIM-SM hello messages

Command	Explanation
Interface Configuration Mode	
ip pim hello-interval <*interval*> **no ip pim hello-interval**	To configure the interval for PIM-SM hello messages. The no form of this command restores the interval to the default value.

Configure the hold time for PIM-SM hello messages

Command	Explanation
Interface Configuration Mode	
ip pim hello-holdtime <*value*> **no ip pim hello-holdtime**	To configure the value of the holdtime field in the PIM-SM hello messages. The no form of this command will restore the hold time to the default value.

Configure ACL for PIM-SM neighbors

Command	Explanation
Interface Configuration Mode	
ip pim neighbor-filter{<*access-list-number*> } **no ip pim neighbor-filter{<*access-list-number*> }**	To configure ACL to filter PIM-SM neighbors. If session to the neighbor has been denied by ACL, then the sessions that have been set up will be discarded immediately and new sessions will not be set up.

Configure the interface as the boundary interface of the PIM-SM protocol

Command	Explanation
Interface Configuration Mode	
ip pim bsr-border **no ip pim bsr-border**	To configure the interface as the boundary of PIM-SM protocol. On the boundary interface, BSR messages will not be sent or received. The network connected the interface is considered as directly connected network. The no form of this command will remove the configuration.

Configure the interface as the management boundary of the PIM-SM protocol

Command	Explanation
Interface Configuration Mode	
ip pim scope-border <*1-99* > | **<*acl_name*>** **no ip pim scope-border**	To configure PIM-SM management boundary for the interface and apply ACL for the management boundary. With default settings, 239.0.0.0/8 is considered as the scope of the management group. If ACL is configured, then the scope specified by ACL permit command is the scope of the management group. acl_name should be standard IPv4 ACL name. The no form of this command will remove the configuration.

Configure global PIM-SM parameter

1) Configure the switch as a candidate BSR

Command	Explanation
Global Configuration Mode	
ip pim bsr-candidate {vlan <*vlan-id*>| **<*ifname*>}[<*mask-length*>][<*priority*>]** **no ip pim bsr-candidate**	This command is the global candidate BSR configuration command, which is used to configure the information of PIM-SM candidate BSR so that it can compete for BSR router with other candidate BSR. The "**no ip pim bsr-candidate**" command cancels the configuration of BSR.

2) Configure the switch as a candidate RP

Command	Explanation
Global Configuration Mode	
ip pim rp-candidate { vlan <*vlan-id*>| **lookback<*index*> <*ifname*>}**	This command is the global candidate RP configuration command, which is used to configure

[<*A.B.C.D*>][<*priority*>] **no ip pim rp-candiate**	the information of PIM-SM candidate RP so that it can compete for RP router with other candidate RP. The "**no ip pim rp-candidate**" command cancels the configuration of RP.

3) Configure static RP

Command	Explanation
Global Configuration Mode	
ip pim rp-address <*A.B.C.D*> **[<*A.B.C.D/M*>]** **no ip pim rp-address <*A.B.C.D*>** **{<*all*>|<*A.B.C.D/M*>}**	The command is the multicast group configuration static RP of the globally or multicast address range. The no form of this command will remove the configuration for the static RP.

4) Configure the cache time of kernel multicast route

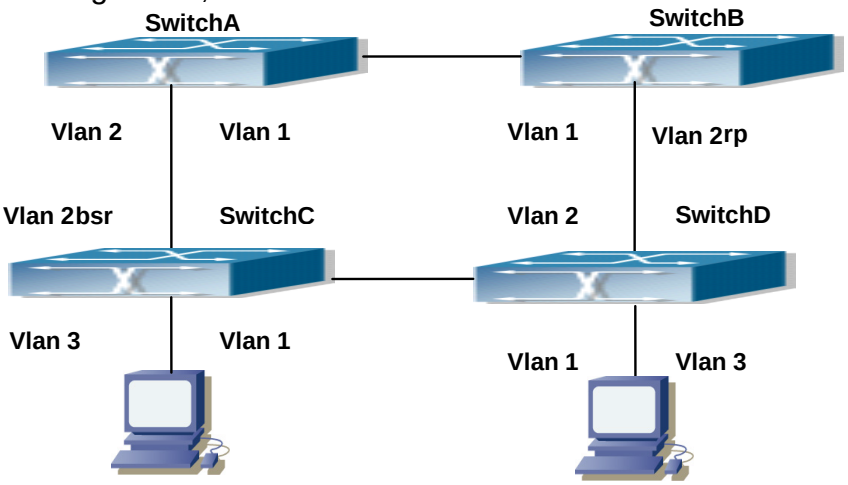
Command	Explanation
Global Configuration Mode	
ip multicast unresolved-cache aging-time <value> **no ip multicast unresolved-cache aging-time**	Configure the cache time of kernel multicast route, the no command restores the default value.

4. Disable PIM-SM Protocol

Command	Explanation
Interface Configuration Mode	
no ip pim sparse-mode | no ip pim multicast-routing(Global configuration mode)	To disable the PIM-SM protocol.

28.3.3 PIM-SM Configuration Examples

As shown in the following figure, add the Ethernet interfaces of SwitchA, SwitchB, SwitchC and SwitchD to corresponding VLAN, and enable PIM-SM Protocol on each VLAN interface.



PIM-SM Typical Environment

The configuration procedure for SwitchA, SwitchB, SwitchC and SwitchD is as follows:

(1) Configure SwitchA:

```
Switch(config)#ip pim multicast-routing
Switch(config)#interface vlan 1
Switch(Config-if-Vlan1)# ip address 12.1.1.1 255.255.255.0
Switch(Config-if-Vlan1)# ip pim sparse-mode
Switch(Config-if-Vlan1)#exit
Switch(config)#interface vlan 2
Switch(Config-if-Vlan2)# ip address 13.1.1.1 255.255.255.0
Switch(Config-if-Vlan2)# ip pim sparse-mode
```

(2) Configure SwitchB:

```
Switch(config)#ip pim multicast-routing
Switch(config)#interface vlan 1
Switch(Config-if-Vlan1)# ip address 12.1.1.2 255.255.255.0
Switch(Config-if-Vlan1)# ip pim sparse-mode
Switch(Config-if-Vlan1)#exit
Switch(config)#interface vlan 2
Switch(Config-if-Vlan2)# ip address 24.1.1.2 255.255.255.0
Switch(Config-if-Vlan2)# ip pim sparse-mode
Switch(Config-if-Vlan2)# exit
Switch(config)# ip pim rp-candidate vlan2
```

(3) Configure SwitchC:

```
Switch(config)#ip pim multicast-routing
Switch(config)#interface vlan 1
Switch(Config-if-Vlan1)# ip address 34.1.1.3 255.255.255.0
Switch(Config-if-Vlan1)# ip pim sparse-mode
```

Switch(Config-if-Vlan1)#exit
Switch(config)#interface vlan 2
Switch(Config-if-Vlan2)# ip address 13.1.1.3 255.255.255.0
Switch(Config-if-Vlan2)# ip pim sparse-mode
Switch(Config-if-Vlan2)#exit
Switch(config)#interface vlan 3
Switch(Config-if-Vlan3)# ip address 30.1.1.1 255.255.255.0
Switch(Config-if-Vlan3)# ip pim sparse-mode
Switch(Config-if-Vlan3)# exit
Switch(config)# ip pim bsr-candidate vlan2 30 10
(4) Configure SwitchD:
Switch(config)#ip pim multicast-routing
Switch(config)#interface vlan 1
Switch(Config-if-Vlan1)# ip address 34.1.1.4 255.255.255.0
Switch(Config-if-Vlan1)# ip pim sparse-mode
Switch(Config-if-Vlan1)#exit
Switch(config)#interface vlan 2
Switch(Config-if-Vlan2)# ip address 24.1.1.4 255.255.255.0
Switch(Config-if-Vlan2)# ip pim sparse-mode
Switch(Config-if-Vlan2)#exit
Switch(config)#interface vlan 3
Switch(Config-if-Vlan3)# ip address 40.1.1.1 255.255.255.0
Switch(Config-if-Vlan3)# ip pim sparse-mode
At the same time, you should pay attention to the configuration of Unicast Routing Protocol, assure that each device can communicate with each other in the network layer, and be able to implement dynamic routing update in virtue of Unicast Routing Protocol.

28.3.4 PIM-SM Troubleshooting

In configuring and using PIM-SM Protocol, PIM-SM Protocol might not operate normally caused by physical connection or incorrect configuration. Therefore, the user should pay attention to the following issues:
Assure that physical connection is correct;
Assure the Protocol of Interface and Link is UP (use show interface command);
Assure that PIM Protocol is enabled in Global Mode (use ip pim multicast-routing);
Assure that PIM-SM is configured on the interface (use ip pim sparse-mode);
Multicast Protocol requires RPF Check using unicast routing; therefore the correctness of unicast routing must be assured beforehand;
PIM-SM Protocol requires supports by RP and BSR, therefore you should use show ip pim bsr-router first to see if there is BSR information. If not, you need to check if there is unicast

routing leading to BSR.

Use show ip pim rp-hash command to check if RP information is correct; if there is not RP information, you still need to check unicast routing.

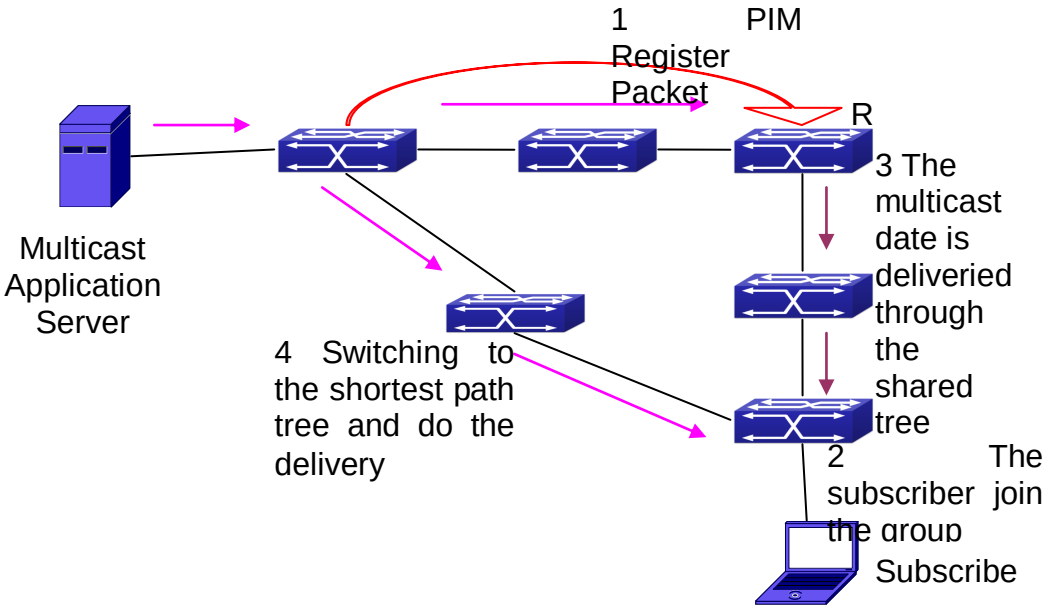
If all attempts including Check are made but the problems on PIM-SM can't be solved yet, then use debug commands such debug pim/debug pim BSR please, and then copy DEBUG information in 3 minutes and send to Technology Service Center.

28.4 MSDP Configuration

28.4.1 Introduction to MSDP

MSDP – Multicast Source Discovery Protocol, is a protocol that can learn information about multicast source in other PIM-SM domain. The RP on which MSDP is configured will advertise the information about the multicast sources in its domain to all the other MSDP entities through SA messages. Thus, all the information about multicast sources in one PIM-SM domain is spread to another. In MSDP, inter-domain information tree is used other than the shared tree. It is required that the multicast routing protocol used for in-domain routing must be PIM-SM.

The work flow for RP in PIM-SM protocol



28.4.2 Brief Introduction to MSDP Configuration Tasks

Configuration of MSDP Basic Function

Enabling MSDP (Required)

Configuring MSDP entities (Required)

Configuring the Connect-Source interface

Configuring static RPF entities

Configuring Originator RP
Configuring TTL value
Configuration of MSDP entities
Configuring the Connect-Source interface
Configuring the descriptive information for MSDP entities
Configuring the AS number
Configuring the specified mesh group of MSDP
Configuring the maximum size for the cache
Configurations on delivery of SA packets
Configuring filter policies for creation of SA packets
Configuring filter rules on how to receive and forward SA packets
Configuring SA request packets
Configuring filter policies for SA-Request packets
Configuration of parameters of SA-cache
Configuring SA packets cache
Configuring the aging time for entries in SA packets cache
Configuring the maximum size for the cache

28.4.3 Configuration of MSDP Basic Function

All the commands in this section are configured for RP in the PIM-SM domain. These RP will function as the other peer of the MSDP entities.

28.4.3.1 Prerequisites of MSDP Configuration

Before the MSDP basic functions can be configured, the following tasks should be done:
At least one single cast routing protocol should be configured, in order to connect the network inside the domain and outside
Configure PIM-SM in order to implement multicast inside the domain

When configuring MSDP basic function, the following information should be ready:
The IP address of MSDP entities
Filter policy table
Pay attention: MSDP can not use with Any-cast RP at same time, but configure Any-cast RP of based MSDP protocol.

28.4.3.2 Enabling MSDP

MSDP should be enabled before various MSDP functions can be configured.
Enable the MSDP function
Configure MSDP

1. Enabling MSDP

Commands	Explanation
Global Configuration Mode	
router msdp **no router msdp**	To enable MSDP. The no form of this command will disable MSDP globally.

2. Configuration of MSDP parameters

Commands	Explanation
MSDP Configuration Mode	
connect-source <interface-type> <interface-number> **no connect-source**	To configure the Connect-Source interface for MSDP Peer. The no form of this command will remove the configured Connect-Source interface.
default-rpf-peer <peer-address> [rp-policy <acl-list-number>|<word>] **no default-rpf-peer**	To configure static RPF Peer. The no form of this command will remove the configured RPF Peer.
originating-rp <interface-type> <interface-number> **no originating-rp**	To configure Originator-RP. The no form of this command will remove the configured Originator-RP.
ttl-threshold <*ttl*> **no ttl-threshold**	To configure the TTL value. The no form of this command will remove the configured TTL value.

28.4.4 Configuration of MSDP Entities

28.4.4.1 Creation of MSDP Peer

Commands	Explanation
MSDP Configuration Mode	
peer <peer-address> **no peer <peer-address>**	To create a MSDP Peer. The no form of this command will remove the configured MSDP Peer.

28.4.4.2 Configuration of MSDP parameters

Commands	Explanation
MSDP Peer Configuration Mode	
connect-source <interface-type>	To configure the Connect-Source interface for

<interface-number> **no connect-source**	MSDP Peer. The no form of this command will remove the configured Connect-Source interface.
description <*text*> **no description**	To configure the descriptive information about the MSDP entities. The no form of this command will remove the configured description.
remote-as <*as-num*> **no remote-as <*as-num*>**	To configure the AS number for MSDP Peer. The no form of this command will remove the configured AS number of MSDP Peer.
mesh-group <*name*> **no mesh-group <*name*>**	To configure an MSDP Peer to join the specified mesh group. The no form of this command will remove the MSDP Peer from the specified mesh group.

28.4.5 Configuration of Delivery of MSDP Packet

Commands	Explanation
MSDP Configuration Mode	
redistribute [list <acl-list-number |acl-name>] **no redistribute**	To configure the filter rules for creation of SA packets. The no form of this command will remove the configured.
MSDP Configuration Mode or MSDP Peer Configuration Mode	
sa-filter(in|out) [list <acl-number | acl-name> | rp-list <rp-acl-number | rp-acl-name>] **no sa-filter(in|out) [[list <acl-number | acl-name> | rp-list <rp-acl-number | rp-acl-name>]**	To configure the filter rules for receiving and forwarding SA packets. The no form of this command will remove the configured rules.
MSDP Peer Configuration Mode	
sa-request **no sa-request**	To configure sending of SA request packets. The no form of this command will disable sending of SA request packets.
MSDP Configuration Mode	
sa-request-filter [list <access-list-number | access-list-name>] **no sa-request-filter [list <access-list-**	To configure filter rules for receiving SA request packets. The no form of this command will remove the configured filter

number | access-list-name>]	rules for SA request packets.

28.4.6 Configuration of Parameters of SA-cache

Commands	Explanation
MSDP Configuration Mode	
cache-sa-state **no cache-sa-state**	To enable the SA packet cache. To disable the SA packets cache.
MSDP Configuration Mode	
cache-sa-holdtime <*150-3600*> **no cache-sa-holdtime**	The aging time for entries in the SA cache. To restore the default aging time configuration.
MSDP Configuration Mode or MSDP Peer Configuration Mode	
cache-sa-maximum <*sa-limit*> **no cache-sa-maximum**	To configure the maximum size for the SA cache. To restore the size of the SA cache to the default value.

28.4.7 MSDP Configuration Examples

Example 1: MSDP basic function.

Multicast Configuration:

Suppose the multicast server is sending multicast datagram at 224.1.1.1;

The designated router – DR, which is connected to the multicast server, encapsulate the multicast datagram in the Register packets and send them to the RP(RP1) in the local domain;

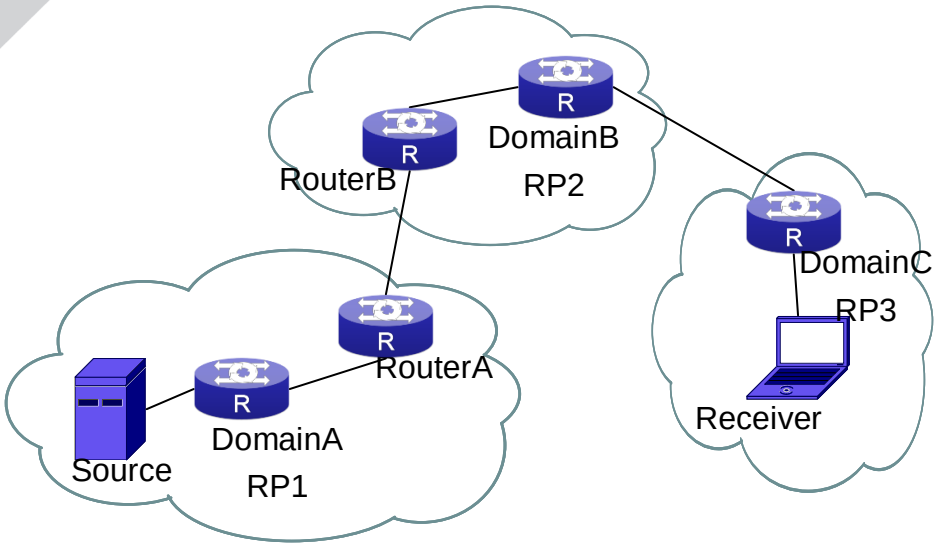
The RP unwraps the packets and sends them to all the domain members through the shared tree. The members in the domain can be configured to be or not to be in the shared tree;

At the same time, the source RP in the domain, generates a SA – Source Active message, and send it to the MSDP entity – RP2.

If there's another member in the same domain with the MSDP entity which is named as RP3, RP3 will distribute the multicast datagram encapsulated in the SA messages to the members of the shared tree, and send join messages to the multicast source. That means RP creates an entry (S, G), and send join messages for (S, G) hop by hop, so that (S, G) can reach the SPT which takes the multicast source as the root across the PIM-SM domain.

If there no members in the same domain with MSDP entity – RP2, RP2 will not create the (S, G) entry nor it will join the SPT which takes the multicast source as the root.

When the reverse route has been set up, the multicast datagram from the source will be directly delivered to RP3, and RP will forward the datagram to the shared tree. At this time, the router which is closest to the domain members can determine itself whether or not to switch to SPT.



Network Topology for MSDP Entry

Configuration tasks are listed as below:

Prerequisites:

Enable the single cast routing protocol and PIM protocol on every router, and make sure that the inter-domain routing works well and multicasting inside the domain works well.

Suppose the multicast server S in Domain A offers multicast programs at 224.1.1.1. A host in Domain C named R subscribes this program. Before MSDP is configured C cannot subscribe the multicast program. However, with the following configuration, R is able to receive programs offered by S.

RP1 in Domain A:

```
Switch#config
Switch(config)#interface vlan 1
Switch(Config-if-Vlan1)#ip address 10.1.1.1 255.255.255.0
Switch(Config-if-Vlan1)#exit
Switch(config)#router msdp
Switch(router-msdp)#peer 10.1.1.2
```

Router A in Domain A:

```
Switch#config
Switch(config)#interface vlan 1
Switch(Config-if-Vlan1)#ip address 10.1.1.2  255.255.255.0
Switch(Config-if-Vlan1)#exit
Switch(config)#interface vlan 2
Switch(Config-if-Vlan2)#ip address 20.1.1.2  255.255.255.0
Switch(Config-if-Vlan2)#exit
Switch(config)#router msdp
```

Switch(router-msdp)#peer 10.1.1.1
Switch(msdp-peer)#exit
Switch(router-msdp)#peer 20.1.1.1

Router B in Domain B:
Switch#config
Switch(config)#interface vlan 2
Switch(Config-if-Vlan2)#ip address 20.1.1.1 255.255.255.0
Switch(Config-if-Vlan2)#exit
Switch(Config)#interface vlan 3
Switch(Config-if-Vlan3)#ip address 30.1.1.1 255.255.255.0
Switch(Config-if-Vlan3)#exit
Switch(config)#router msdp
Switch(router-msdp)#peer 20.1.1.2
Switch(msdp-peer)#exit
Switch(router-msdp)#peer 30.1.1.2

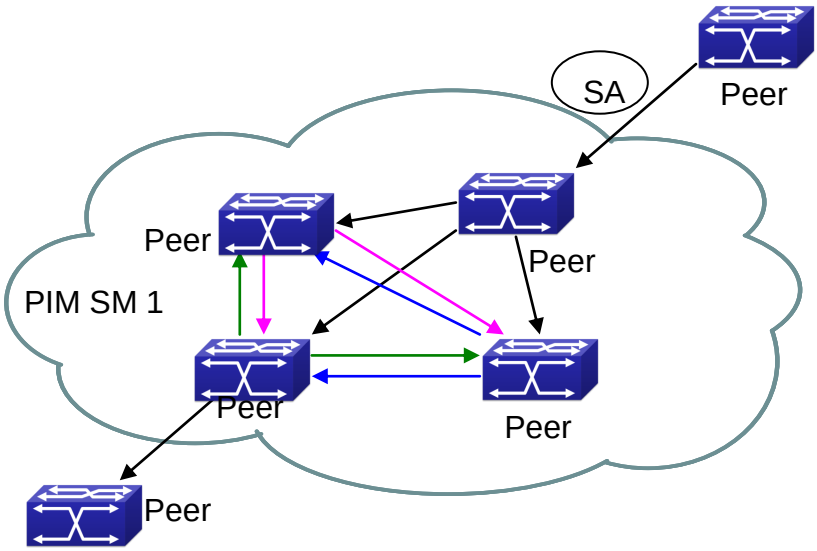
RP2 in Domain B:
Switch#config
Switch(config)#interface vlan 3
Switch(Config-if-Vlan3)#ip address 30.1.1.2 255.255.255.0
Switch(config)#interface vlan 4
Switch(Config-if-Vlan4)#ip address 40.1.1.2 255.255.255.0
Switch(Config-if-Vlan4)#exit
Switch(config)#router msdp
Switch(router-msdp)#peer 30.1.1.1
Switch(config)#router msdp
Switch(router-msdp)#peer 40.1.1.1

RP3 in Domain C:
Switch(config)#interface vlan 4
Switch(Config-if-Vlan1)#ip address 40.1.1.1 255.255.255.0
Switch(Config-if-Vlan1)#exit
Switch(config)#router msdp
Switch(router-msdp)#peer 40.1.1.2

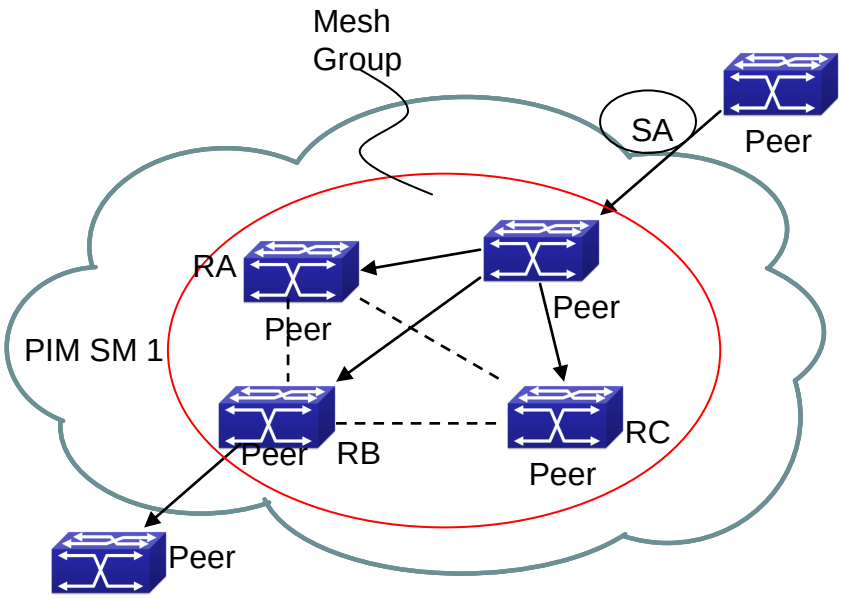
Example 2: Application of MSDP Mesh-Group.
Mesh-Group can be used to reduce flooding of SA messages. The Peers which are meshed in the same domain can be configured as a Mesh-Group. All the members in the same mesh

group use a unique group name.

As it is shown in Figure, when Mesh-Group is configured for the four meshed Peers in the same domain, flooding of SA messages reduced remarkably.



Flooding of SA messages



Flooding of SA messages with mesh group configuration

Configuration steps are listed as below:

Router A:

Switch#config

Switch(config)#interface vlan 1

Switch(Config-if-Vlan1)#ip address 10.1.1.1 255.255.255.0

Switch(Config-if-Vlan1)#exit

Switch(config)#interface vlan 2
Switch(Config-if-Vlan2)#ip address 20.1.1.1 255.255.255.0
Switch(Config-if-Vlan2)#exit
Switch(config)#interface vlan 3
Switch(Config-if-Vlan3)#ip address 30.1.1.1 255.255.255.0
Switch(Config-if-Vlan3)#exit
Switch(config)#router msdp
Switch(router-msdp)#peer 10.1.1.2
Switch(router-msdp)#mesh-group QTECH-1
Switch(msdp-peer)#exit
Switch(router-msdp)#peer 20.1.1.4
Switch(router-msdp)#mesh-group QTECH-1
Switch(msdp-peer)#exit
Switch(router-msdp)#peer 30.1.1.3
Switch(router-msdp)#mesh-group QTECH-1
Switch(msdp-peer)#exit

Router B:
Switch#config
Switch(config)#interface vlan 1
Switch(Config-if-Vlan1)#ip address 10.1.1.2 255.255.255.0
Switch(Config-if-Vlan1)#exit
Switch(config)#interface vlan 4
Switch(Config-if-Vlan4)#ip address 40.1.1.2 255.255.255.0
Switch(Config-if-Vlan4)#exit
Switch(config)#interface vlan 6
Switch(Config-if-Vlan6)#ip address 60.1.1.2 255.255.255.0
Switch(Config-if-Vlan6)#exit
Switch(config)#router msdp
Switch(router-msdp)#peer 10.1.1.1
Switch(router-msdp)#mesh-group QTECH-1
Switch(msdp-peer)#exit
Switch(router-msdp)#peer 40.1.1.4
Switch(router-msdp)#mesh-group QTECH -1
Switch(msdp-peer)#exit
Switch(router-msdp)#peer 60.1.1.3
Switch(router-msdp)#mesh-group QTECH -1

Router C:

```
Switch#config
Switch(config)#interface vlan 4
Switch(Config-if-Vlan4)#ip address 40.1.1.4  255.255.255.0
Switch(Config-if-Vlan4)#exit
Switch(config)#interface vlan 5
Switch(Config-if-Vlan5)#ip address 50.1.1.4  255.255.255.0
Switch(Config-if-Vlan5)#exit
Switch(config)#interface vlan 6
Switch(Config-if-Vlan6)#ip address 60.1.1.4  255.255.255.0
Switch(Config-if-Vlan6)#exit
Switch(config)#router msdp
Switch(router-msdp)#peer 20.1.1.1
Switch(router-msdp)#mesh-group QTECH -1
Switch(msdp-peer)#exit
Switch(router-msdp)#peer 40.1.1.4
Switch(router-msdp)#mesh-group QTECH -1
Switch(msdp-peer)#exit
Switch(router-msdp)#peer 60.1.1.2
Switch(router-msdp)#mesh-group QTECH -1
```

Router D:

```
Switch#config
Switch(config)#interface vlan 2
Switch(Config-if-Vlan2)#ip address 20.1.1.4 255.255.255.0
Switch(Config-if-Vlan2)#exit
Switch(config)#interface vlan 4
Switch(Config-if-Vlan1)#ip address 40.1.1.4  255.255.255.0
Switch(Config-if-Vlan1)#exit
Switch(config)#interface vlan 5
Switch(Config-if-Vlan5)#ip address 50.1.1.4  255.255.255.0
Switch(Config-if-Vlan5)#exit
Switch(config)#router msdp
Switch(router-msdp)#peer 20.1.1.1
Switch(router-msdp)#mesh-group QTECH -1
Switch(msdp-peer)#exit
Switch(router-msdp)#peer 40.1.1.2
Switch(router-msdp)#mesh-group QTECH -1
Switch(msdp-peer)#exit
Switch(router-msdp)#peer 50.1.1.3
```

Switch(router-msdp)#mesh-group QTECH -1

28.4.8 MSDP Troubleshooting

When MSDP is being configured, it may not function because of the physical link not working or configuration mistakes. Attention should be paid to the following items in order to make MSDP work:

Make sure the physical link works well

Make sure inner-domain and inter-domain routing works

Make sure PIM-SM is applied in every domain as the inner-domain routing protocol, and configuration for PIM-SM works well

Make sure MSDP is enabled, and the link status of the MSDP enabled Peer is UP

Use the command **show msdp global** to check whether the MSDP configuration is correct

If the MSDP problems cannot be solved through all the methods provided above, please issue the command **debug msdp** to get the debugging messages within three minutes, and send them to the technical service center of our company.

28.5 ANYCAST RP Configuration

28.5.1 Introduction to ANYCAST RP

Anycast RP is a technology based on PIM protocol, which provides redundancy in order to recover as soon as possible once an RP becomes unusable.

The kernel concept of Anycast RP is that the RP addresses configured all over the whole network exist on multiple multicast servers (the most common situation is that every device providing ANYCAST RP uses LOOPBACK interface, and using the longest mask to configures RP addresses on this interface), while the unicast routing algorithm will make sure that PIM routers can always find the nearest RP, thus , providing a shorter and faster way to find RP in a larger network., Once an RP being used becomes unusable, the unicast routing algorithm will ensure that the PIM router can find a new RP path fast enough to recover the multicast server in time. Multiple RP will cause a new problem that is if the multicast source and the receivers are registered to different RP, some receivers will not be able to receive data of multicast source (obviously, the register messages only prefer the nearest RP). So, in order to keep the communication between all RP, Anycast RP defines that the nearest RP to the multicast source should forward the source register messages to all the other RP to guarantee that all joiners of the RP can find the multicast source.

The method to realize the PIM-protocol-based Anycast RP is that: maintaining an ANYCAST RP list on every switch configured with Anycast RP and using another address as the label to identify each other. When one Anycast RP device receives a register message, it will send the register message to other Anycast RP devices while using its own address as the source

address, to notify all the other devices of the original destination.

28.5.2 ANYCAST RP Configuration Task

1. Enable ANYCAST RP v4 function
2. Configure ANYCAST RP v4

1. Enable ANYCAST RP v4 function

Command	Explanation
Global Configuration Mode	
ip pim anycast-rp **no ip pim anycast-rp**	Enable ANYCAST RP function. (necessary) No operation will globally disable ANYCAST RP function.

2. Configure ANYCAST RP v4

(1) Configure the RP candidate

Command	Explanation
Global Configuration Mode	
ip pim rp-candidate {vlan<*vlan-id*> **|loopback<*index*> |<*ifname*>} [<*A.B.C.D*>]** **[<*priority*>]** **no ip pim rp-candidate**	Now, the PIM-SM has allowed the Loopback interface to be a RP candidate.(necessary) Please pay attention to that, ANYCAST RP protocol can configure the Loopback interface or a regular three-layer VLAN interface to be the RP candidate. In make sure that PIM routers in the network can find where the RP locates, the RP candidate interface should be added into the router. No operation will cancel the RP candidate configuration on this router.

(2) Configure self-rp-address (the RP address of this router)

Command	Explanation
Global Configuration Mode	
ip pim anycast-rp self-rp-address A.B.C.D **no ip pim anycast-rp self-rp-address**	Configure the self-rp-address of this router (as a RP). This address can be used to exclusively identify this router when communicating with other RP. the effect of **self-rp-address** refers to two respects: 1 Once this router (as a RP) receives the

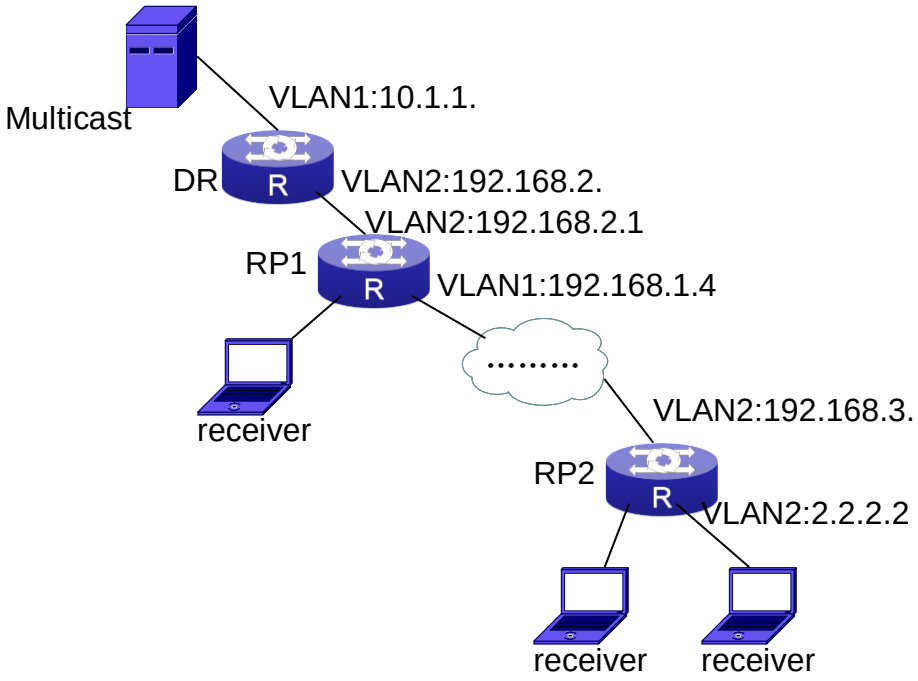
	register message from DR unicast, it needs to forward the register message to all the other RP in the network, notifying them of the state of source (S.G). While forwarding the register message, this router will change the source address of it into self-rp-address. 2 Once this router(as a RP) receives a register message from other RP unicast, such as a register message whose destination is the self-rp-address of this router, it will create (S,G) state and send back a register-stop message, whose destination address is the source address of the register message. Pay attention: self-rp-address has to be the address of a three-layer interface on this router, but the configuration is allowed to be done with the absence of the interface. The self-rp-address should be unique. No operation will cancel the self-rp-address which is used to communicate with other RPs by this router (as a RP).

(3) Configure other-rp-address (other RP communication addresses)

Command	Explanation
Global Configuration Mode	
ip pim anycast-rp <anycast-rp-addr> <other-rp-addr> **no ip pim anycast-rp <anycast-rp-addr> <other-rp-addr>**	Configure anycast-rp-addr on this router (as a RP). This unicast address is actually the RP address configured on multiple RP in the network, in accordance with the address of RP candidate interface (or Loopback interface). The effect of **anycast-rp-addr** includes: 1 Although more than one anycast-rp-addr addresses are allowed to be configured, only the one having the same address with the currently configured RP candidate address will take effect. Only after that, can the other-rp-address in accordance with this anycast-rp-addr take effect.

	2 The configuration is allowed to be done with the absence of the interface in accordance with the anycast-rp-addr. Configure on this router (as a RP) the other-rp-addresses of other RP communicating with it. This unicast address identifies other RP and is used in the communication with local routers. The effect of **other-rp-address** refers to two respects: 1 Once this router (as a RP) receives the register message from a DR unicast, it should forward it to other RP in the network to notify all the RP in the network of the source (S.G) state. While forwarding, the router will change the destination address of the register message into other-rp-address. 2 Multiple other-rp-addresses can be configured in accordance with one anycast-rp-addr, Once the register message from a DR is received, it should be forwarded to all of these other RP one by one. No operation will cancel an other-rp-address communicating with this router.

28.5.3 ANYCAST RP Configuration Examples



The ANYCAST RP v4 function of the router

As shown in the Figure, the overall network environment is PIM-SM, which provides two routers supporting ANYCAST RP, RP1 and RP2. Once multicast data from the multicast source server reaches the DR, the DR will send a multicast source register message to the nearest RP unicast according to the unicast routing algorithm, which is RP1 in this example. When RP1 receives the register message from the DR, besides redistributing to the shared tree according to the orderers who already join it, it will forward the multicast register message to RP2 to guarantee that all orders that already join RP2 can find the multicast source. Since there is an ANYCAST list maintained on router RP1 that has been configured with ANYCAST RP, and since this list contains the unicast addresses of all the other RP in the network, when the RP1 receives the register message, it can use the self-r-address, which identifies itself as the source address to forward the register message to RP2. The cloud in the Figure represents the PIM-SM network operation between RP1 and RP2.

The following is the configuration steps:
RP1 Configuration:
Switch#config
Switch(config)#interface loopback 1
Switch(Config-if-Loopback1)#ip address 1.1.1.1 255.255.255.255
Switch(Config-if-Loopback1)#exit
Switch(config)#ip pim rp-candidate loopback1
Switch(config)#ip pim bsr-candidate vlan 1
Switch(config)#ip pim multicast-routing

Switch(config)#ip pim anycast-rp
Switch(config)#ip pim anycast-rp self-rp-address 192.168.2.1
Switch(config)#ip pim anycast-rp 1.1.1.1 192.168.3.2
RP2 Configuration:
Switch#config
Switch(config)#interface loopback 1
Switch(Config-if-Loopback1)#ip address 1.1.1.1 255.255.255.255
Switch(Config-if-Loopback1)#exit
Switch(config)#ip pim rp-candidate loopback1
Switch(config)#ip pim multicast-routing
Switch(config)#ip pim anycast-rp
Switch(config)#ip pim anycast-rp self-rp-address 192.168.3.2
Switch(config)#ip pim anycast-rp 1.1.1.1 192.168.2.1

28.5.4 ANYCAST RP Troubleshooting

When configuring and using ANYCAST RP function, the ANYCAST RP might work abnormally because of faults in physical connections, configurations or something others. So, the users should pay attention to the following points:
The physical connections should be guaranteed to be correct
The PIM-SM protocol should be guaranteed to operate normally
The ANYCAST RP should be guaranteed to be enabled in Global configuration mode
The self-rp-address should be guaranteed to be configured correctly in Global configuration mode
The other-rp-address should be guaranteed to be configured correctly in Global configuration mode
All the interface routers should be guaranteed to be correctly added, including the loopback interface as a RP
Use “**show ip pim anycast rp status**” command to check whether the configuration information of ANYCAST RP is correct
If the problems of ANYCAST still cannot be solved after checking, please use debug commands like “**debug pim anycast-rp**”, then copy the DEBUG information within three minutes and send it to the technical service center of our company.

28.6 PIM-SSM

28.6.1 Introduction to PIM-SSM

Source Specific Multicast (PIM-SSM) is a new kind of multicast service protocol. With PIM-SSM, a multicast session is distinguished by the multicast group address and multicast source

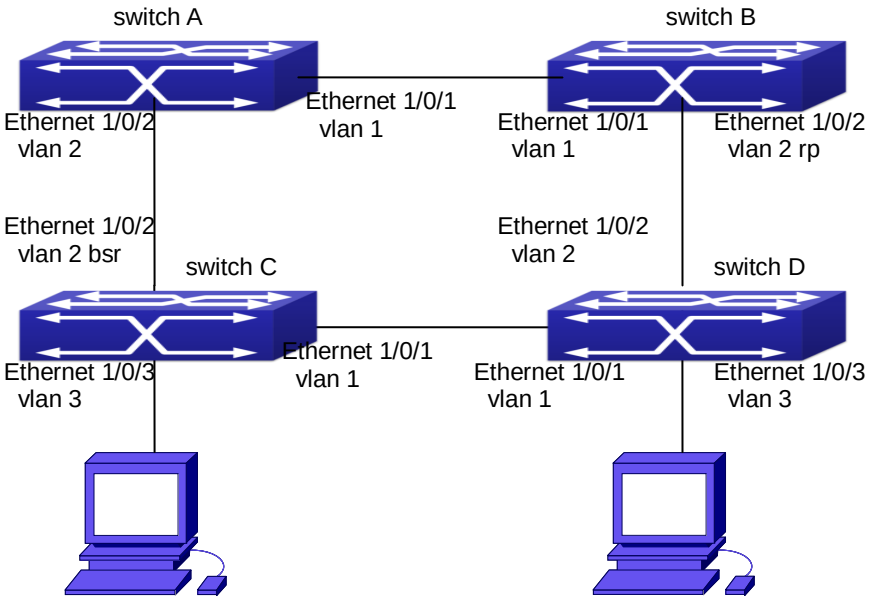
address. In SSM, hosts can be added into the multicast group manually and efficiently like the traditional PIM-SM, but leave out the shared tree and RP management in PIM-SM. In SSM, SPT tree will be constructed with (S, G). G for the multicast group address and S for the source address of the multicast which sends datagram to G. (S, G) in a pair is named as a channel of SSM. SSM serves best for the application of multicast service which is from one station to many ones, for example, the network sports video channel, and the news channel. By default, the multicast group address of SSM is limited between 232.0.0.0 and 232.255.255.255. However this address range can be extended according to actual situations.

28.6.2 PIM-SSM Configuration Task List

Command	Explanation
Global Configuration Mode	
ip multicast ssm {default|range *<access-list-number >*} **no ip multicast ssm**	To configure the address range for pim-ssm. The no form command will disable the configuration.

28.6.3 PIM-SSM Configuration Examples

As the figure shows, ethernet interfaces from SwitchA, SwitchB, SwitchC, and SwitchD are configured to be in separate VLANs. And PIM-SSM is enabled globally by enabling the PIM-SM or PIM-DM protocol on the VLAN interfaces. Take PIM-SM for example.



PIM-SSM typical environment

Configurations of SwitchA, SwitchB, SwitchC, and SwitchD are shown as below.

(1) Configuration of SwitchA.

Switch(config)#ip pim multicast-routing
Switch(config)#interface vlan 1
Switch(Config-If-Vlan1)# ip pim sparse-mode
Switch(Config-If-Vlan1)#exit
Switch(config)#interface vlan 2
Switch(Config-If-Vlan2)# ip pim sparse-mode
Switch(Config-If-Vlan2)#exit
Switch(config)#access-list 1 permit 224.1.1.1 0.0.0.255
Switch(config)#ip multicast ssm range 1
(2) Configuration of SwitchB.
Switch(config)#ip pim multicast-routing
Switch(config)#interface vlan 1
Switch(Config-If-Vlan1)# ip pim sparse-mode
Switch(Config-If-Vlan1)#exit
Switch(config)#interface vlan 2
Switch(Config-If-Vlan2)# ip pim sparse-mode
Switch(Config-If-Vlan2)# exit
Switch(config)# ip pim rp-candidate vlan2
Switch(config)#access-list 1 permit 224.1.1.1 0.0.0.255
Switch(config)#ip multicast ssm range 1
(3) Configuration of SwitchC.
Switch(config)#ip pim multicast-routing
Switch(config)#interface vlan 1
Switch(Config-If-Vlan1)# ip pim sparse-mode
Switch(Config-If-Vlan1)#exit
Switch(config)#interface vlan 2
Switch(Config-If-Vlan2)# ip pim sparse-mode
Switch(Config-If-Vlan2)#exit
Switch(config)#interface vlan 3
Switch(Config-If-Vlan3)# ip pim sparse-mode
Switch(Config-If-Vlan3)# exit
Switch(config)# ip pim bsr-candidate vlan2 30 10
Switch(config)#access-list 1 permit 224.1.1.1 0.0.0.255
Switch(config)#ip multicast ssm range 1
(4) Configuration of SwitchD.
Switch(config)#ip pim multicast-routing
Switch(config)#interface vlan 1
Switch(Config-If-Vlan1)# ip pim sparse-mode
Switch(Config-If-Vlan1)#exit

Switch(config)#interface vlan 2
Switch(Config-If-Vlan2)# ip pim sparse-mode
Switch(Config-If-Vlan2)#exit
Switch(config)#interface vlan 3
Switch(Config-If-Vlan3)# ip pim sparse-mode
Switch(Config-If-Vlan3)#exit
Switch(config)#access-list 1 permit 224.1.1.1 0.0.0.255
Switch(config)#ip multicast ssm range 1

28.6.4 PIM-SSM Troubleshooting

In configuring and using PIM-SSM Protocol, PIM-SSM Protocol might not operate normally caused by physical connection or incorrect configuration. Therefore, the user should pay attention to the following issues:
Assure that physical connection is correct;
Assure the Protocol of Interface and Link is UP (use **show interface** command);
Assure that PIM Protocol is enabled in Global Mode (use **ip pim multicast-routing**);
Assure that PIM-SSM is configured on the interface (use **ip pim sparse-mode**);
Assure that SSM is configured in Global Mode;
Multicast Protocol requires RPF check using unicast routing, therefore the correctness of unicast routing must be assured beforehand.
If all attempts including check are made but the problems on PIM-SSM can't be solved yet, then use debug commands such **debug pim event/debug pim packet** please, and then copy DEBUG information in 3 minutes and send to Technology Service Center.

28.7 DVMRP

28.7.1 Introduction to DVMRP

DVMRP Protocol, namely, is "Distance Vector Multicast Routing Protocol". It is a Multicast Routing Protocol in dense mode, which sets up a Forward Broadcast Tree for each source in a manner similar to RIP, and sets up a Truncation Broadcast Tree, i.e. the Shortest Path Tree to the source, for each source through dynamic Prune/Graft.
Some of the important features of DVMRP are:
The routing exchange used to determine reverse path checking information is based on distance vector (in a manner similar to RIP)
Routing exchange update occurs periodically (the default is 60 seconds)
TTL upper limit = 32 hops (and that RIP is 16)
Routing update includes net mask and supports CIDR
In comparison with Unicast routing, Multicast routing is a kind of reverse routing (that is, what

you are interested in is where the packets are from but not where they go), thus the information in DVMRP routing table is used to determine if an input Multicast packet is received at the correct interface. Otherwise, the packet will be discarded to prevent Multicast circulation.

The check which determines if the packet gets to the correct interface is called RPF check. When some Multicast data packets get to some interface, it will determine the reverse path to the source network by looking up DVMRP router table. If the interface data packets get to is the one which is used to send Unicast message to the source, then the reverse path check is correct, and the data packets are forwarded out from all downstream interfaces. If not, then probably there is failure, and the Multicast packet is discarded.

Since not all switches support Multicast, DVMRP supports tunnel multicast communication, tunnel is a method to send multicast data report among DVMRP switches separated by switches which don't support multicast routing. Multicast data packets are encapsulated in unicast data packets and directly sent to the next switch which supports multicast. DVMRP Protocol treats tunnel interface and general physical interface equally.

If two or more switches are connected to a multi-entrance network, it is likely to transmit more than one copy of a data packet to the sub-network. Thus a specified transmitter must be appointed. DVMRP achieves this goal by making use of routing exchange mechanism; when two switches on the multi-entrance network exchange routing information, they will be aware of the routing distance from each other to the source network, thus the switch with the shortest distance to the source network will become the specified transmitter of the sub-network. If some have the same distance, then the one with the lowest IP prevails.

After some interface of the switch is configured to Function DVMRP Protocol, the switch will multicast Probe message to other DVMRP switches on this interface, which is used to find neighbors and detect the capabilities of each other. If no Probe message from the neighbor is received until the neighbor is timed out, then this neighbor is considered missing.

In DVMRP, source network routing selection message are exchanged in a basic manner same to RIP. That is, routing report message is transmitted among DVMRP neighbors periodically (the default is 60 seconds). The routing information in DVMRP routing selection table is used to set up source distribution tree, i.e. to determine by which neighbor it passes to get to the source transmitting multicast packet; the interface to this neighbor is called upstream interface. The routing report includes source network (use net mask) address and the hop entry for routing scale.

In order to finish transmission correctly, every DVMRP switch needs to know which downstream switches need to receive multicast packet from some specific source network through it. After receiving packets from some specific source, DVMRP switch firstly will broadcast these multicast packets from all downstream interfaces, i.e. the interfaces on which there are other DVMRP switches which have dependence on the specific source. After receiving Prune message from some downstream switch on the interface, it will prune this

switch. DVMRP switch makes use of poison reverse to notify the upstream switch for some specific source: “I am your downstream.” By adding infinity (32) to the routing distance of some specific source it broadcasts, DVMRP switch responds to the source upstream exchange to fulfill poison reverse. This means distance correct value is 1 to 2* infinity (32) -1 or 1 to 63, 1 to 63 means it can get to source network, 32 means source network is not arrival, 33 to 63 means the switch which generates the report message will receive multicast packets from specific source depending on upstream router.

28.7.2 DVMRP Configuration Task List

Globally enable and disable DVMRP (Required)
Configure Enable and Disable DVMRP Protocol at the interface (Required)
Configure DVMRP Sub-parameters (Optional)
Configure DVMRP interface parameters
Configure the delay of transmitting report message on DVMRP interface and the message number each time it transmits
Configure metric value of DVMRP interface
Configure if DVMRP is able to set up neighbors with DVMRP routers which can not Prune/Graft
Configure DVMRP tunnel

1. Globally enable DVMRP Protocol

The basic configuration to function DVMRP routing protocol on QTECH series Layer 3 switch is very simple. Firstly it is required to turn on DVMRP switch globally.

Command	Explanation
Global Mode	
[no] ip dvmrp multicast-routing	Globally enable DVMRP Protocol, the “**no ip dvmrp multicast-routing**” command disables DVMRP Protocol globally. (Required)

2. Enable DVMRP Protocol on the interface

The basic configuration to function DVMRP routing protocol on QTECH series Layer 3 switch is very simple. After globally enabling DVMRP Protocol, it is required to turn on DVMRP switch under corresponding interface.

Command	Explanation
Interface Configuration Mode	
ip dvmrp **no ip dvmrp**	Enable DVMRP Protocol on the interface, the “**no ip dvmrp**” command disables DVMRP Protocol on the interface.

3. Configure DVMRP Sub-parameters

(1) Configure DVMRP Interface Parameters

Configure the delay of transmitting report message on DVMRP interface and the message number each time it transmits

Configure metric value of DVMRP interface

Configure if DVMRP is able to set up neighbors with DVMRP routers which can not Prune/Graft

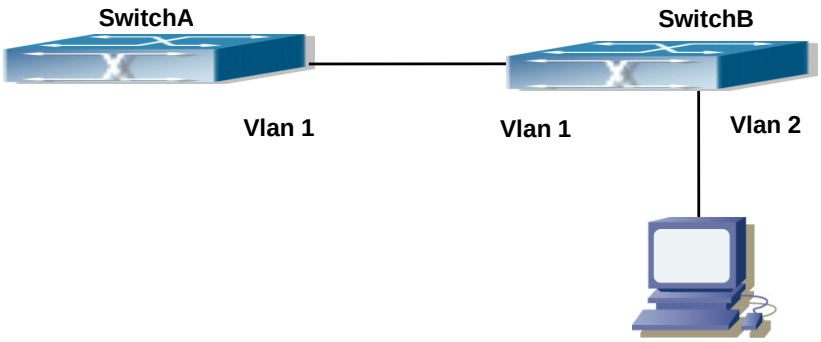
Command	Explanation
Interface Configuration Mode	
ip dvmrp output-report-delay <*delay_val*> [<*burst_size*>] **no ip dvmrp output-report-delay**	Configure the delay of transmitting DVMRP report message on interface and the message number each time it transmits, the "**no ip dvmrp output-report-delay**" command restores default value.
ip dvmrp metric <*metric_val*> **no ip dvmrp metric**	Configure interface DVMRP report message metric value; the "**no ip dvmrp metric**" command restores default value.
ip dvmrp reject-non-pruners **no ip dvmrp reject-non-pruners**	Configure the interface rejects to set up neighbor relationship with non pruning/grafting DVMRP router. The "**no ip dvmrp reject-non-pruners**" command restores to being able to set up neighbor ship.

4. Configure DVMRP Tunnel

Command	Explanation
Interface Configuration Mode	
ip dvmrp tunnel <index> <src-ip> <dst-ip> **no ip dvmrp tunnel {<index> |<src-ip> <dst-ip>}**	This command configures a DVMRP tunnel; the "**no ip dvmrp tunnel {<*index*> |<*src-ip*> <*dst-ip*>}**" command deletes a DVMRP tunnel.

28.7.3 DVMRP Configuration Examples

As shown in the following figure, add the Ethernet interfaces of Switch A and Switch B to corresponding VLAN, and enable DVMRP on each VLAN interface.



DVMRP Network Topology Diagram

The configuration procedure for SwitchA and SwitchB is as follows:
(1) Configure SwitchA:
Switch (config)#ip dvmrp multicast-routing
Switch (config)#interface vlan 1
Switch(Config-if-Vlan1)# ip address 10.1.1.1 255.255.255.0
Switch(Config-if-Vlan1)# ip dvmrp enable
(2) Configure SwitchB:
Switch (config)#ip dvmrp multicast-routing
Switch (config)#interface vlan 1
Switch(Config-if-Vlan1)# ip address 12.1.1.2 255.255.255.0
Switch(Config-if-Vlan1)# ip dvmrp enable
Switch(Config-if-Vlan1)#exit
Switch (config)#interface vlan 2
Switch(Config-if-Vlan2)# ip address 20.1.1.1 255.255.255.0
Switch(Config-if-Vlan2)# ip dvmrp
Since DVMRP itself does not rely on Unicast Routing Protocol, it is not necessary to configure Unicast Routing Protocol. This is the difference from PIM-DM and PIM-SM.

28.7.4 DVMRP Troubleshooting

In configuring and using DVMRP Protocol, DVMRP Protocol might not operate normally caused by physical connection or incorrect configuration. Therefore, the user should pay attention to the following issues:
Firstly to assure that physical connection is correct;
Next, to assure the Protocol of Interface and Link is UP (use **show interface** command);
Please check if the correct IP address is configured on the interface (use **ip address** command);
Afterwards, enable DVMRP Protocol on the interface (use **ip dvmrp** command and **ip dv multicast-routing** command);
Multicast Protocol requires RPF Check using unicast routing; therefore the correctness of

unicast routing must be assured beforehand. (DVMRP uses its own unicast table, please use **show ip dvmrp route** command to look up).
If all attempts including Check are made but the problems on DVMRP can't be solved yet, then please use commands such as debug DVMRP, and then copy DEBUG information in 3 minutes and send to Technology Service Center.

28.8 DCSCM

28.8.1 Introduction to DCSCM

DCSCM (Destination control and source control multicast) technology mainly includes three aspects, i.e. Multicast Packet Source Controllable, Multicast User Controllable and Service-Oriented Priority Strategy Multicast.
The Multicast Packet Source Controllable technology of Security Controllable Multicast technology is mainly processed in the following manners:
On the edge switch, if source under-control multicast is configured, then only multicast data from specified group of specified source can pass.
For RP switch in the core of PIM-SM, for REGISTER information out of specified source and specified group, REGISTER_STOP is transmitted directly and table entry is not allowed to set up. (This task is implemented in PIM-SM model).
The implement of Multicast User Controllable technology of Security Controllable Multicast technology is based on the control over IGMP report message sent out by the user, thus the model being controlled is IGMP snooping and IGMPmodel, of which the control logic includes the following three, i.e. to take control based on VLAN+MAC address transmitting packets, to take control based on IP address of transmitting packets and to take control based on the port where messages enter, in which IGMP snooping can use the above three methods to take control simultaneously, while since IGMP model is located at layer 3, it only takes control over the IP address transmitting packets.
The Service-Oriented Priority Strategy Multicast of Security Controllable technology adopts the following mode: for multicast data in limit range, set the priority specified by the user at the join-in end so that data can be sent in a higher priority on TRUNK port, consequently guarantee the transmission is processed in user-specified priority in the entire network.

28.8.2 DCSCM Configuration Task List

Source Control Configuration
Destination Control Configuration
Multicast Strategy Configuration

Source Control Configuration

Source Control Configuration has three parts, of which the first is to enable source control. The command of source control is as follows:

Command	Explanation
Global Configuration Mode	
[no] ip multicast source-control (Required)	Enable source control globally, the "**no ip multicast source-control**" command disables source control globally. It is noticeable that, after enabling source control globally, all multicast packets are discarded by default. All source control configuration can not be processed until that it is enabled globally, while source control can not be disabled until all configured rules are disabled.

The next is to configure the rule of source control. It is configured in the same manner as for ACL, and uses ACL number of 5000-5099, every rule number can be used to configure 10 rules. It is noticeable that these rules are ordered, the front one is the one which is configured the earliest. Once the configured rules are matched, the following rules won't take effect, so rules of globally allow must be put at the end. The commands are as follows:

Command	Explanation
Global Configuration Mode	
[no] access-list <*5000-5099*> {deny|permit} ip {{<*source*> <*source-wildcard*>}|{host-source <*source-host-ip*>}|any-source} {{<*destination*> <*destination-wildcard*>}|{host-destination <destination-host-ip>}|any-destination}	The rule used to configure source control. This rule does not take effect until it is applied to specified port. Using the NO form of it can delete specified rule.

The last is to configure the configured rule to specified port.

Note: If the rules being configured will occupy the table entries of hardware, configuring too many rules will result in configuration failure caused by bottom table entries being full, so we suggest user to use the simplest rules if possible. The configuration rules are as follows:

Command	Explanation
Port Configuration Mode	
[no] ip multicast source-control access-group <*5000-5099*>	Used to configure the rules source control uses to port, the NO form cancels the configuration.

Destination Control Configuration

Like source control configuration, destination control configuration also has three steps.

First, enable destination control globally. Since destination control need to prevent unauthorized user from receiving multicast data, the switch won't broadcast the multicast data

it received after configuring global destination control. Therefore, It should be avoided to connect two or more other Layer 3 switches in the same VLAN on a switch on which destination control is enabled. The configuration commands are as follows:

Command	Explanation
Global Configuration Mode	
[no] multicast destination-control (required)	Globally enable IPv4 and IPv6destination control. The no operation of this command will globally disable destination control. All of the other configuration can only take effect after globally enabled. The next is configuring destination control rules, which are similar.

Next is to configure destination control rule. It is similar to source control, except to use ACL No. of 6000-7999.

Command	Explanation
Global Configuration Mode	
[no] access-list <*6000-7999*> {deny|permit} ip {{<*source*> <*source-wildcard*>}|{host-source <*source-host-ip*>}|any-source} {{<*destination*> <*destination-wildcard*>}|{host-destination <*destination-host-ip*>}|any-destination}	The rule used to configure destination control. This rule does not take effect until it is applied to source IP or VLAN-MAC and port. Using the NO form of it can delete specified rule.

The last is to configure the rule to specified source IP, source VLAN MAC or specified port. It is noticeable that, due to the above situations, these rules can only be used globally in enabling IGMP-SNOOPING. And if IGMP-SNOOPING is not enabled, then only source IP rule can be used under IGMP Protocol. The configuration commands are as follows:

Command	Explanation
Port Configuration Mode	
[no] ip multicast destination-control access-group <*6000-7999*>	Used to configure the rules destination control uses to port, the NO form cancels the configuration.
Global Configuration Mode	
[no] ip multicast destination-control <*1-4094*> <*macaddr*> access-group <6000-7999>	Used to configure the rules destination control uses to specify VLAN-MAC, the NO form cancels the configuration.
[no] ip multicast destination-control <*IPADDRESS/M*> access-group <*6000-7999*>	Used to configure the rules destination control uses to specified IP address/net mask, the NO form cancels the configuration.

Multicast Strategy Configuration

Multicast Strategy uses the manner of specifying priority for specified multicast data to achieve and guarantee the effects the specific user requires. It is noticeable that multicast data can not get a special care all along unless the data are transmitted at TRUNK port. The configuration is very simple, it has only one command, i.e. to set priority for the specified multicast. The commands are as follows:

Command	Explanation
Global Configuration Mode	
[no] ip multicast policy <*IPADDRESS/M*> <*IPADDRESS/M*> cos <*priority*>	Configure multicast strategy, specify priority for sources and groups in specific range, and the range is <0-7>.

28.8.3 DCSCM Configuration Examples

Source Control

In order to prevent an Edge Switch from putting out multicast data ad asbitsium, we configure Edge Switch so that only the switch at port Ethernet1/0/5 is allowed to transmit multicast, and the data group must be 225.1.2.3. Also, switch connected up to port Ethernet1/0/10 can transmit multicast data without any limit, and we can make the following configuration.

EC(config)#access-list 5000 permit ip any host 225.1.2.3

EC(config)#access-list 5001 permit ip any any

EC(config)#ip multicast source-control

EC(config)#interface ethernet1/0/5

EC(Config-If-Ethernet1/0/5)#ip multicast source-control access-group 5000

EC(config)#interface ethernet1/0/10

EC(Config-If-Ethernet1/0/10)#ip multicast source-control access-group 5001

Destination Control

We want to limit users with address in 10.0.0.0/8 network segment from entering the group of 238.0.0.0/8, so we can make the following configuration:

Firstly enable IGMP snooping in the VLAN it is located (Here it is assumed to be in VLAN2)

EC(config)#ip igmp snooping

EC(config)#ip igmp snooping vlan 2

After that, configure relative destination control access-list, and configure specified IP address to use that access-list.

Switch(config)#access-list 6000 deny ip any 238.0.0.0 0.255.255.255

Switch(config)#access-list 6000 permit ip any any

Switch(config)#multicast destination-control

Switch(config)#ip multicast destination-control 10.0.0.0/8 access-group 6000

In this way, users of this network segment can only join groups other than 238.0.0.0/8.

Multicast strategy

Server 210.1.1.1 is distributing important multicast data on group 239.1.2.3, we can configure on its join-in switch as follows:

Switch(config)#ip multicast policy 210.1.1.1/32 239.1.2.3/32 cos 4

In this way, the multicast stream will have a priority of value 4 (Usually this is pretty higher, the higher possible one is protocol data; if higher priority is set, when there is too many multicast data, it might cause abnormal behavior of the switch protocol) when it gets to other switches through this switch.

28.8.4 DCSCM Troubleshooting

The effect of DCSCM module itself is similar to ACL, and the problems occurred are usually related to improper configuration. Please read the descriptions above carefully. If you still can not determine the cause of the problem, please send your configurations and the effects you expect to the after-sale service staff of our company.

28.9 IGMP

28.9.1 Introduction to IGMP

IGMP (Internet Group Management Protocol) is the protocol in TCP/IP protocol family which is responsible for IP multicast member management. It is used to set up and maintain multicast group member relationship between IP host and its neighbor multicast switches. IGMP does not include the spread and maintenance of relation information of group members among multicast switches, this work is accomplished by each multicast routing protocol. All hosts participating in multicast must implement IGMP protocol.

Hosts participating IP multicast can join in and exit multicast group at any location, any time and without limit of member total. Multicast switch does not need and not likely to save all relationships of all hosts. It only gets to know if there are receivers of some multicast group, i.e. group member, on the network segment each interface connects to. And the host only needs to save which multicast groups it joined.

IGMP is asymmetric between host and router: the host needs to respond the IGMP query messages of multicast switches, i.e. to report message response in membership; the switch sends out membership query messages periodically, and then determine if there are hosts of some specific group joining in the sub-network it belongs to based on the received response message, and send out query of specific group (IGMP version2) when receiving the report of a host exiting the group to determine if there exists no member in some specific group.

Up to now, there are three versions of IGMP: IGMP version1 (defined by RFC1112), IGMP version2 (defined by RFC2236) and IGMP version3 (defined by RFC3376).

The main improvements of IGMP version2 over version1 are:

1. The election mechanism of multicast switches on the shared network segment

Shared network segment is the situation of there is more than one multicast switch on a network segment. Under this kind of situation, since all switches which runs IGMP under this network segment can get membership report message from the host, therefore, only one switch is required to transmit membership query message, so an exchange election mechanism is required to determine a switch as query machine. In IGMP version1, the selection of query machine is determined by Multicast Routing Protocol; IGMP version2 made an improvement for it, it prescribed that when there are more than one multicast switches on the same network segment, the multicast switch with the lowest IP address will be elected as the query machine.

2. IGMP version2 added Leave Group Mechanism

In IGMP version 1, the host leaves the multicast group silently without sending any notification to any multicast switch. This causes that the multicast switch can only determine the leave of multicast member by multicast group response time-out. But in version2, when a host decides to leave a multicast group, if it is the host which gives response to the latest membership query message, then it will send out a message implying it is leaving.

3. IGMP version 2 added the query to specific group

In IGMP version1, a query of multicast switch is for all multicast groups on the network segment. This query is called general group query. In IGMP version2, query of specific group is added besides general group query. The destination IP address of this kind of query message is the IP address of the multicast group, the group address field part of the message is also the IP address of the multicast group. Thus it is prevented that hosts which are other multicast group members transmit response message.

4. IGMP version2 added the biggest response time field

IGMP version2 added the biggest response time field to dynamically adjust the response time of the host to group query message.

The main features of version3 is allowing the host to choose receiving from or rejecting a certain source, which is the basis of SSM (Source-Specific Multicast) multicast. For example, when a host is sending a report of INCLUDE{10.1.1.1, 10.1.1.2} to some group G, that means the host needs the router to forward the flux from 10.1.1.1 and 10.1.1.2; when a host is sending a report of EXCLUDE{192.168.1.1} to some group G, that means the host needs the flux from all sources of group G except 192.168.1.1. This makes a great difference from the previous IGMP.

The main improvements of IGMP Version3 over IGMP Version1 and Version2 are:

The status to be maintained is group and source list, not only the groups in IGMPv2.

The interoperations with IGMPv1 and IGMPv2 are defined in IGMPv3 status.

IP service interface is modified to allow specific source list thereby.

The queried includes his/her Robustness Variable and Query Interval in query group to allow

the synchronization with these variables of non-queries.
Max Response Time in Query Message has an exponential range, with maximum value from 25.5 secs of v2 to 53 mins, which can be used in links of great capacity.
In order to increase strength, the host retransmits State-Change message.
Additional data is defined to adapt future extension.
Report group is sent to 224.0.0.22 to help with IGMP Snooping of Layer 2 Switch.
Report group can include more than one group record, and it allows using small group to report complete current status.
The host does not restrain operation any more, which simplifies the implement and allows direct membership trace.
In querying messages, the new router side restraint process (S sign) modified the existing strength of IGMPv2.

28.9.2 IGMP Configuration Task List

Enable IGMP (Required)
Configure IGMP sub-parameters (Optional)
(1) Configure IGMP group parameters
Configure IGMP group filtering conditions
Configure IGMP to join in group
Configure IGMP to join in static group
(2) Configure IGMP query parameters
Configure the interval of IGMP sending query message
Configure the maximum response time of IGMP query
Configure time-out of IGMP query
(3) Configure IGMP version
Disable IGMP Protocol

Enable IGMP Protocol
There are not specific commands for enabling IGMP Protocol on the Layer 3 switch. Enabling any multicast protocol under corresponding interface will automatically enable IGMP.

Command	Explanation
Global Mode	
ip dvmrp multicast-routing | ip pim multicast-routing	To enable global multicast protocol is the prerequisite to enable IGMP protocol, the "**no ip dvmrp multicast-routing | no ip pim multicast-routing**" commands disable multicast protocol and IGMP protocol. (Required)

Command	Explanation

Interface Configuration Mode	
ip dvmrp enable| ip pim dense-mode | ip pim sparse-mode	Enable IGMP Protocol, the corresponding commands "**no ip dvmrp enable| no ip pim dense-mode | no ip pim sparse-mode**" disable IGMP Protocol. (Required)

Configure IGMP Sub-parameters

(1) Configure IGMP group parameters

Configure IGMP group filtering conditions

Configure IGMP to join in group

Configure IGMP to join in static group

Command	Explanation
Interface Configuration Mode	
ip igmp access-group {<*acl_num* | *acl_name*>} **no ip igmp access-group**	Configure the filtering conditions of the interface to IGMP group; the "**no ip igmp access-group**" command cancels the filtering condition.
ip igmp join-group <*A.B.C.D* > **no ip igmp join-group <*A.B.C.D* >**	Configure the interface to join in some IGMP group, the "**no ip igmp join-group <*A.B.C.D* >**" command cancels the join.
ip igmp static-group <*A.B.C.D* > **no ip igmp static-group <*A.B.C.D*>**	Configure the interface to join in some IGMP static group; the "**no ip igmp static-group <*A.B.C.D* >**" command cancels the join.

(2) Configure IGMP Query parameters

Configure interval for IGMP to send query messages

Configure the maximum response time of IGMP query

Configure the time-out of IGMP query

Command	Explanation
Interface Configuration Mode	
ip igmp query-interval <*time_val*> **no ip igmp query-interval**	Configure the interval of IGMP query messages sent periodically; the "**no ip igmp query-interval**" command restores default value.
ip igmp query-max-response-time <*time_val*> **no ip igmp query-max-response-time**	Configure the maximum response time of the interface for IGMP query; the "**no ip igmp query-max-response-time**" command restores default value.
ip igmp query-timeout <*time_val*> **no ip igmp query-timeout**	Configure the time-out of the interface for IGMP query; the "**no ip igmp query-timeout**" command

	restores default value.

(3) Config IGMP version

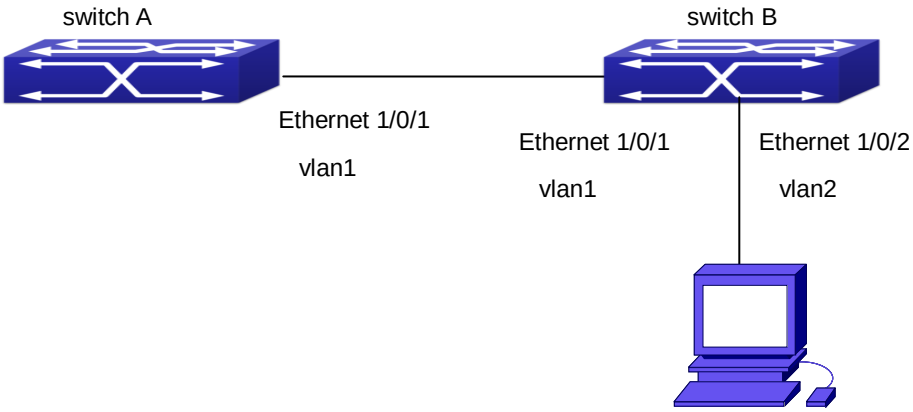
Command	Explanation
Global Mode	
ip igmp version <*version*> **no ip igmp version**	Configure IGMP version on the interface; the "**no ip igmp version**" command restores the default value.

Disable IGMP Protocol

Command	Explanation
Interface Configuration Mode	
no ip dvmrp | no ip pim dense-mode | no ip pim sparse-mode | no ip dvmrp multicast-routing | no ip pim multicast-routing	Disable IGMP Protocol.

28.9.3 IGMP Configuration Examples

As shown in the following figure, add the Ethernet ports of Switch A and Switch B to corresponding VLAN, and start PIM-DM on each VLAN interface.



IGMP Network Topology Diagram

The configuration procedure for SwitchA and SwitchB is as follows:

(1) Configure SwitchA:

Switch(config)#ip pim multicast-routing

Switch (config)#interface vlan 1

Switch(Config-if-Vlan1)#ip address 12.1.1.1 255.255.255.0

Switch(Config-if-Vlan1)#ip pim dense-mode

(2) Configure SwitchB:

```
Switch(config)#ip pim multicast-routing
Switch(config)#interface vlan1
Switch(Config-if-Vlan1)#ip address 12.1.1.2 255.255.255.0
Switch(Config-if-Vlan1)#ip pim dense-mode
Switch(Config-if-Vlan1)#exit
Switch(config)#interface vlan2
Switch(Config-if-Vlan1)#ip address 20.1.1.1 255.255.255.0
Switch(Config-if-Vlan2)#ip pim dense-mode
Switch(Config-if-Vlan2)#ip igmp version 3
```

28.9.4 IGMP Troubleshooting

In configuring and using IGMP Protocol, IGMP Protocol might not operate normally caused by physical connection or incorrect configuration. Therefore, user should pay attention to the following issues:

Firstly to assure that physical connection is correct;

Next, to assure the Protocol of Interface and Link protocol is UP (use show interface command);

Afterwards, to assure to start a kind of multicast protocol on the interface;

Multicast Protocol requires RPF Check using unicast routing; therefore the correctness of unicast routing must be assured beforehand.

28.10 IGMP Snooping

28.10.1 Introduction to IGMP Snooping

IGMP (Internet Group Management Protocol) is a protocol used in IP multicast. IGMP is used by multicast enabled network device (such as a router) for host membership query, and by hosts that are joining a multicast group to inform the router to accept packets of a certain multicast address. All those operations are done through IGMP message exchange. The router will use a multicast address (224.0.0.1) that can address to all hosts to send an IGMP host membership query message. If a host wants to join a multicast group, it will reply to the multicast address of that a multicast group with an IGMP host membership reports a message.

IGMP Snooping is also referred to as IGMP listening. The switch prevents multicast traffic from flooding through IGMP Snooping, multicast traffic is forwarded to ports associated to multicast devices only. The switch listens to the IGMP messages between the multicast router and hosts, and maintains multicast group forwarding table based on the listening result, and can then decide to forward multicast packets according to the forwarding table.

Switch provides IGMP Snooping and is able to send a query from the switch so that the user can use switch in IP multicast.

28.10.2 IGMP Snooping Configuration Task List

1. Enable IGMP Snooping
2. Configure IGMP Snooping

1. Enable IGMP Snooping

Command	Explanation
Global Mode	
ip igmp snooping **no ip igmp snooping**	Enables IGMP Snooping. The no operation disables IGMP Snooping function.

2. Configure IGMP Snooping

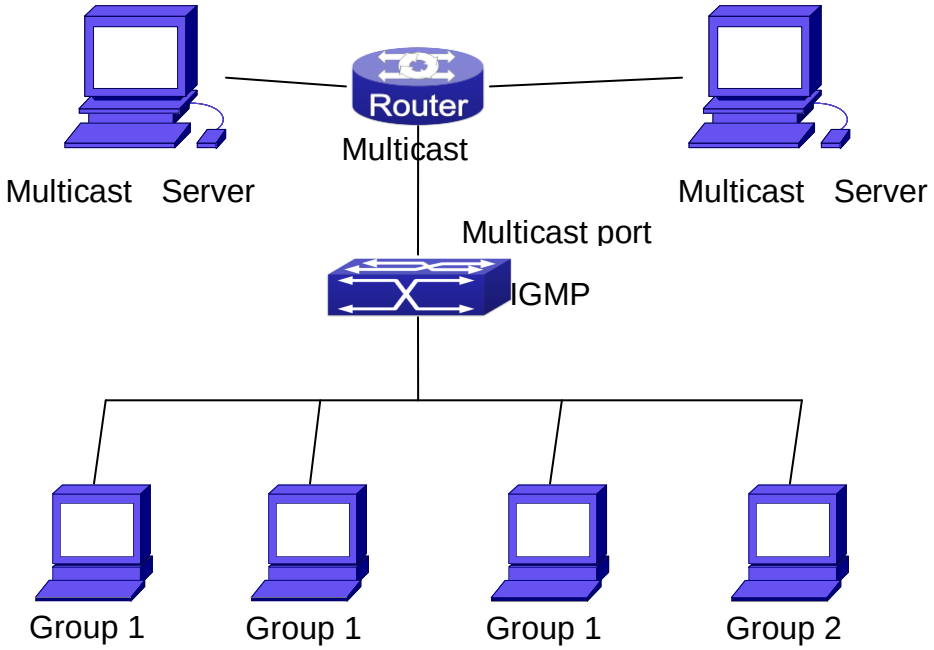
Command	Explanation
Global Mode	
ip igmp snooping vlan <*vlan-id*> **no ip igmp snooping vlan <*vlan-id*>**	Enables IGMP Snooping for specified VLAN. The no operation disables IGMP Snooping for specified VLAN.
ip igmp snooping proxy **no ip igmp snooping proxy**	Enable IGMP Snooping proxy function, the no command disables the function.
ip igmp snooping vlan < *vlan-id* > limit {group <*g_limit*> | source <*s_limit*>} **no ip igmp snooping vlan < *vlan-id* > limit**	Configure the max group count of vlan and the max source count of every group. The "**no ip igmp snooping vlan <*vlan-id*> limit**" command cancels this configuration.
ip igmp snooping vlan <*vlan-id*> l2-general-querier **no ip igmp snooping vlan <*vlan-id*> l2-general-querier**	Set this vlan to layer 2 general querier. It is recommended to configure a layer 2 general querier on a segment. The "**no ip igmp snooping vlan <*vlan-id*> l2-general-querier**"command cancels this configuration.
ip igmp snooping vlan <vlan-id> l2-general-querier-version <version>	Configure the version number of a general query from a layer 2 general querier.
ip igmp snooping vlan <vlan-id> l2-general-querier-source <source>	Configure the source address of a general query from a layer 2 general querier.
ip igmp snooping vlan <*vlan-id*> mrouter-port interface <*interface –name*> **no ip igmp snooping vlan <*vlan-id*> mrouter-port interface <*interface –name*>**	Configure static mrouter port of vlan. The no form of the command cancels this configuration.

ip igmp snooping vlan <*vlan-id*> mrouter-port learnpim **no ip igmp snooping vlan <*vlan-id*> mrouter-port learnpim**	Enable the function that the specified VLAN learns mrouter-port (according to pim packets), the no command will disable the function.
ip igmp snooping vlan <*vlan-id*> mrpt <*value* > **no ip igmp snooping vlan <*vlan-id*> mrpt**	Configure this survive time of mrouter port. The "**no ip igmp snooping vlan <*vlan-id*> mrpt**" command restores the default value.
ip igmp snooping vlan <*vlan-id*> query-interval <*value*> **no ip igmp snooping vlan <*vlan-id*> query-interval**	Configure this query interval. The "**no ip igmp snooping vlan <*vlan-id*> query-interval**" command restores the default value.
ip igmp snooping vlan <*vlan-id*> immediately-leave **no ip igmp snooping vlan <*vlan-id*> immediately-leave**	Enable the IGMP fast leave function for the specified VLAN: the "**no ip igmp snooping vlan <*vlan-id*> immediate-leave**" command disables the IGMP fast leave function.
ip igmp snooping vlan <*vlan-id*> query-mrsp <*value*> **no ip igmp snooping vlan <*vlan-id*> query-mrsp**	Configure the maximum query response period. The "**no ip igmp snooping vlan <*vlan-id*> query-mrsp**" command restores to the default value.
ip igmp snooping vlan <vlan-id> query-robustness <*value*> **no ip igmp snooping vlan <vlan-id> query-robustness**	Configure the query robustness. The "**no ip igmp snooping vlan <vlan-id> query-robustness**" command restores to the default value.
ip igmp snooping vlan <*vlan-id*> suppression-query-time <*value*> **no ip igmp snooping vlan <*vlan-id*> suppression-query-time**	Configure the suppression query time. The "**no ip igmp snooping vlan <*vlan-id*> suppression-query-time**" command restores to the default value.
ip igmp snooping vlan <*vlan-id*> static-group <*A.B.C.D*> [source <*A.B.C.D*>] interface [ethernet | port-channel] <*IFNAME*> **no ip igmp snooping vlan <*vlan-id*> static-group <*A.B.C.D*> [source <*A.B.C.D*>] interface [ethernet | port-channel] <*IFNAME*>**	Configure static-group on specified port of the VLAN. The no form of the command cancels this configuration.

ip igmp snooping vlan <*vlan-id*> report source-address <*A.B.C.D*> **no ip igmp snooping vlan <*vlan-id*> report source-address**	Configure forwarding IGMP packet source address, The no operation cancels the packet source address.
ip igmp snooping vlan <vlan-id> specific-query-mrsp <value> **no ip igmp snooping vlan <vlan-id> specific-query-mrspt**	Configure the maximum query response time of the specific group or source, the no command restores the default value.

28.10.3 IGMP Snooping Examples

Scenario 1: IGMP Snooping function



Enabling IGMP Snooping function

Example: As shown in the above figure, a VLAN 100 is configured in the switch and includes ports 1, 2, 6, 10 and 12. Four hosts are connected to port 2, 6, 10, 12 respectively and the multicast router is connected to port 1. As IGMP Snooping is disabled by default either in the switch or in the VLANs, If IGMP Snooping should be enabled in VLAN 100, the IGMP Snooping should be first enabled for the switch in Global Mode and in VLAN 100 and set port 1 of VLAN 100 to be the mrouter port.

The configuration steps are listed below:
Switch(config)#ip igmp snooping
Switch(config)#ip igmp snooping vlan 100
Switch(config)#ip igmp snooping vlan 100 mrouter interface ethernet 1/0/1

Multicast Configuration

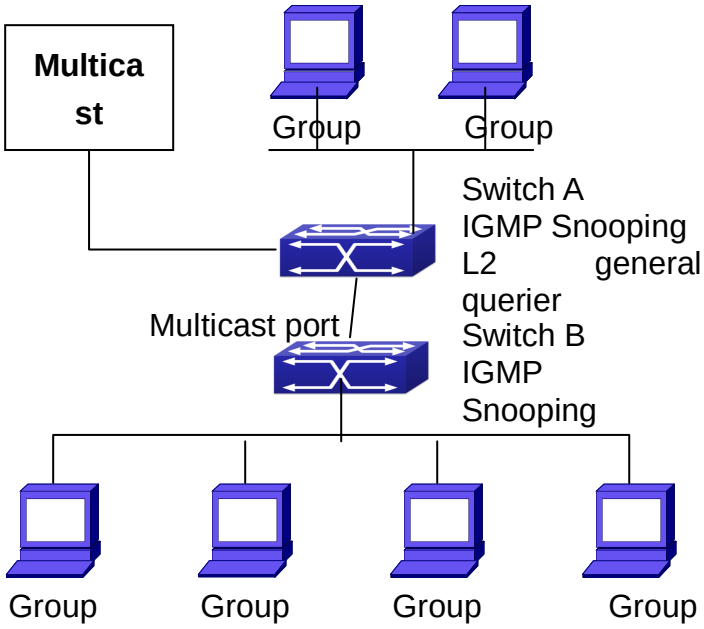
Suppose two programs are provided in the Multicast Server using multicast address Group1 and Group2, three of four hosts running multicast applications are connected to port 2, 6, 10 plays program1, while the host is connected to port 12 plays program 2.

IGMP Snooping listening result:

The multicast table built by IGMP Snooping in VLAN 100 indicates ports 1, 2, 6, 10 in Group1 and ports 1, 12 in Group2.

All the four hosts can receive the program of their choice: ports 2, 6, 10 will not receive the traffic of program 2 and port 12 will not receive the traffic of program 1.

Scenario 2: L2-general-querier



The switches as IGMP Queries

The configuration of Switch2 is the same as the switch in scenario 1, SwitchA takes the place of Multicast Router in scenario 1. Let's assume VLAN 60 is configured in SwitchA, including ports 1, 2, 6, 10 and 12. Port 1 connects to the multicast server, and port 2 connects to Switch2. In order to send Query at regular interval, IGMP query must enabled in Global mode and in VLAN60.

The configuration steps are listed below:

SwitchA#config

SwitchA(config)#ip igmp snooping

SwitchA(config)#ip igmp snooping vlan 60

SwitchA(config)#ip igmp snooping vlan 60 L2-general-querier

SwitchB#config

SwitchB(config)#ip igmp snooping
SwitchB(config)#ip igmp snooping vlan 100
SwitchB(config)#ip igmp snooping vlan 100 mrouter interface ethernet 1/0/1

Multicast Configuration
The same as scenario 1
IGMP Snooping listening result:
Similar to scenario 1

Scenario 3: To run in cooperation with layer 3 multicast protocols.
switch which is used in Scenario 1 is replaced with ROUTER with specific configurations remains the same. And multicast and IGMP snooping configurations are the same with what it is in Scenario 1. To configure PIM-SM on ROUTER, and enable PIM-SM on vlan 100 (use the same PIM mode with the connected multicast router)

Configurations are listed as below:
switch#config
switch(config)#ip pim multicast-routing
switch(config)#interface vlan 100
switch(config-if-vlan100)#ip pim sparse-mode

IGMP snooping does not distribute entries when layer 3 multicast protocol is enabled. It only does the following tasks.
Remove the layer 2 multicast entries.
Provide query functions to the layer 3 with vlan, S, and G as the parameters.
When layer 3 IGMP is disabled, re-enable distributing layer 2 multicast entries.

By looking up the layer 3 IPMC entries, it can be found that ports can be indicated by the layer 3 multicast entries. This ensures the IGMP snooping can work in cooperation with the layer 3 multicast protocols.

28.10.4 IGMP Snooping Troubleshooting

On IGMP Snooping function configuration and usage, IGMP Snooping might not run properly because of physical connection or configuration mistakes. So the users should note that:
Make sure correct physical connection
Activate IGMP Snooping on whole configuration mode (use **ip igmp snooping**)
Configure IGMP Snooping at VLAN on whole configuration mode (use **ip igmp snooping vlan <vlan-id>**)
Make sure one VLAN is configured as L2 common checker in same mask, or make sure

configured static mrouter
Use **show ip igmp snooping vlan <vid>** command check IGMP Snooping information

28.11 IGMP Proxy Configuration

28.11.1 Introduction to IGMP Proxy

IGMP/MLD proxy which is introduced in rfc4605, is a simplified multicast protocol running at edge boxes. The edge boxes which runs the IGMP/MLD proxy protocol, does not need to run complicated multicast routing protocols such as PIM/DVMRP. However they work with multicast protocol enabled network through IGMP/MLD proxy. They can simplify the implementation of multicasting on edge devices.
The IGMP/MLD proxy works between the multicast router and the client, it works as both the multicast host and router. Upstream and downstream ports should be specified in the IGMP/MLD proxy configuration. The host protocol runs at upstream ports, while the router protocol runs at downstream ports. The switch collects the join and leave messages received from downstream ports and forward them to the multicast router through upstream ports.
The IGMP proxy configuration is exclusive with PIM and DVMRP configuration.

28.11.2 IGMP Proxy Configuration Task List

Enable IGMP Proxy function
Enable configurations for both downstream and upstream ports for the IGMP Proxy in different interfaces
Configure IGMP Proxy

1. Enable IGMP Proxy function

Command	Explanation
Global Mode	
ip igmp proxy **no ip igmp proxy**	Enable IGMP Proxy function. The "**no ip igmp proxy**" disables this function.

2. Enable configurations for both downstream and upstream ports for the IGMP Proxy in different interfaces

Command	Explanation
Interface Configuration Mode	
ip igmp proxy upstream **no ip igmp proxy upstream**	Enable IGMP Proxy upstream function. The "**no ip igmp proxy upstream**" disables this function.
ip igmp proxy downstream **no ip igmp proxy downstream**	Enable IGMP Proxy downstream function. The "**no ip igmp proxy downstream**" disables this

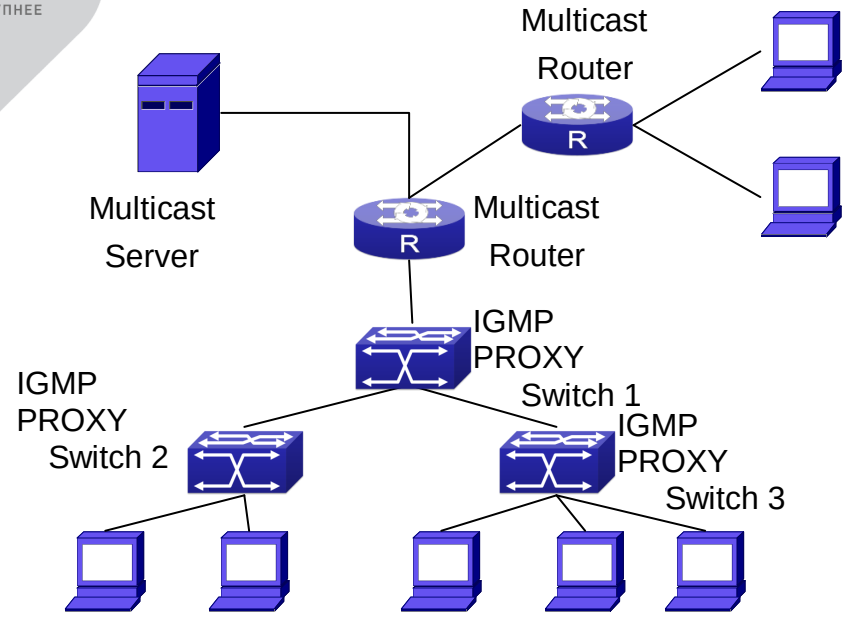
	function.

3. Configure IGMP Proxy assistant parameter

Command	Explanation
Global Mode	
ip igmp proxy limit {group <1-500>| source <1-500>} **no ip igmp proxy limit**	To configure the maximum number of groups that upstream ports can join, and the maximum number of sources in a single group. The no form of this command will restore the default value.
ip igmp proxy unsolicited-report interval <1-5> **no ip igmp proxy unsolicited-report interval**	To configure how often the upstream ports send out unsolicited report. The no form of this command will restore the default configuration.
ip igmp proxy unsolicited-report robustness <2-10> **no ip igmp proxy unsolicited-report robustness**	To configure the retry times of upstream ports' sending unsolicited reports. The no form of this command will restore the default value.
ip igmp proxy aggregate **no ip igmp proxy aggregate**	To configure non-query downstream ports to be able to aggregate the IGMP operations. The no form of this command will restore the default configuration.
ip multicast ssm range <1-99> **ip multicast ssm default** **no ip mulitcast ssm**	To configure the address range for IGMP proxy ssm multicast groups; The no form of this command will remove the configuration.
ip igmp proxy multicast-source **no ip igmp proxy multicast-source**	To configure the port as downstream ports for the source of multicast datagram; The no from of this command will disable the configuration.

28.11.3 IGMP Proxy Examples

Example 1: IGMP Proxy function.



IGMP Proxy Topology Diagram

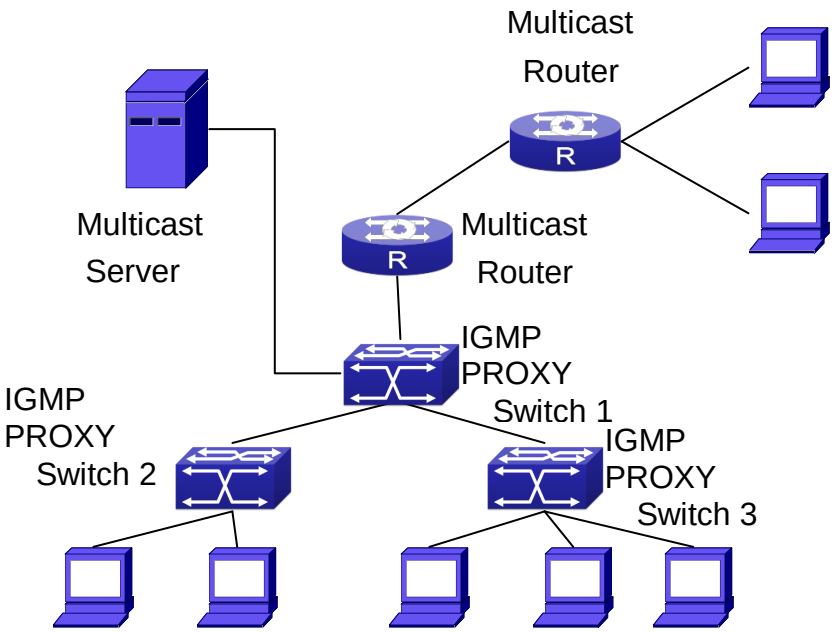
As it is show in the figure above, the switch functions as IGMP Proxy in a network of topology of tree, the switch aggregates the multicast dataflow from upstream port and redistributes them to the downstream ports, while the IGMP membership reports flow from downstream ports to upstream ports. Three IGMP Proxy enabled switches which are connected in tree topology, respectively have one port connected to multicast routers, and no less than one ports connected to hosts or upstream ports from other IGMP Proxy enabled switches.

The configuration steps are listed below:
Switch#config
Switch(config)#ip igmp proxy
Switch(Config)#interface vlan 1
Switch(Config-if-Vlan1)#ip igmp proxy upstream
Switch(Config)#interface vlan 2
Switch(Config-if-Vlan2)#ip igmp proxy downstream

Multicast Configuration:
Suppose the multicast server offers some programs through 224.1.1.1. Some hosts subscribe that program at the edge of the network. The IGMP multicast members report themselves to the downstream ports of IGMP Proxy enabled Switch 2 and Switch 3. Switch 2 and Switch 3 then aggregate the group membership information and send them through the upstream ports. Switch 1 finally forward these membership information to the multicast router when receiving the group membership information through upstream ports, and deliver the multicast dataflow

through downstream ports.

Example2: IGMP Proxy for multicast sources from downstream ports.



IGMP Proxy for multicast sources from downstream ports

As it is show in the figure above, IGMP Proxy enabled switches connected to the network in tree topology. The multicast source server connects to the downstream port of Switch1, the multicast dataflow is distributed through the upstream port and other downstream ports. Three IGMP Proxy enabled switches which are connected in tree topology, respectively have one port connected to multicast routers, and no less than one ports connected to hosts or upstream ports from other IGMP proxy enabled switches.

The configuration steps are listed below:
IGMP PROXY Switch1 configuration:
Switch#config
Switch(config)#ip igmp proxy
Switch(Config)#interface vlan 1
Switch(Config-if-Vlan1)#ip igmp proxy upstream
Switch(Config)#interface vlan 2
Switch(Config-if-Vlan2)#ip igmp proxy downstream
Switch(Config-if-Vlan2)#ip igmp proxy multicast-source
Route1 configuration:
Switch#config
Switch(config)#ip pim multicast
Switch(Config)#interface vlan 1

Switch(Config-if-Vlan1)#ip pim sparse-mode
Switch(Config-if-Vlan1)#ip pim bsr-border

Multicast Configuration:
Suppose the server provides programs through the multicast address 224.1.1.1, and some hosts subscribe that program on the edge of the network. The host reports their IGMP multicast group membership to Switch 2 and Switch 3 through downstream ports. Switch 2 and Switch 3 then aggregate and forward them to Switch 1 which then forwards the information to multicast router. When multicast dataflow arrives, the IGMP Proxy enabled switches re-distribute the group membership through upstream ports and downstream ports. When the multicast router receives the multicast dataflow from IGMP proxy, it will consider the multicast data source is directly connected to the router, and determine the identity of DR and ORIGINATOR. The multicast dataflow will be redistributed according to the PIM protocol.

28.11.4 IGMP Proxy Troubleshooting

When IGMP Proxy function configuration and usage, IGMP Proxy might not run properly because of physical connection or configuration mistakes. So the users should note that:
Make sure physical connection correctly;
Activate IGMP Proxy on whole Global mode (use **ip igmp proxy**);
Make sure configure one upstream port and at least one downstream port under interface configuration mode (Use **ip igmp proxy upstream**, **ip igmp proxy downstream**);
Use **show ip igmp proxy** command to check if the IGMP Proxy information is correct.

If the IGMP Proxy problem remains unsolved, please use debug IGMP Proxy and other debugging command and copy the DEBUG message within three minutes, send the recorded message to the technical service center of our company.

Chapter 29 IPv6 Multicast Protocol

29.1 PIM-DM6

29.1.1 Introduction to PIM-DM6

PIM-DM6(Protocol Independent Multicast, Dense Mode) is the IPv6 version of Protocol Independent Multicast Dense Mode. It is a Multicast Routing Protocol in dense mode which adapted to small network. The members of multicast group are relatively dense under this kind of network environment. There is no difference compared with the IPv4 version PIM-DM except that the addresses it uses are IPv6 addresses. Thus we don't differentiate between PIM-DM and PIM-DM6 in this chapter. All PIM-DM in the text without specific explanation refers to IPv6 version PIM-DM.

As a result of continuous development of IPv6 network, it has the network environment of nonsupport IPv6 multicast sometimes, so it needs to do the IPv6 multicast operation by tunnel. Therefore, our PIM-DM6 supports configuration on configure tunnel, and passes through nonsupport IPv6 multicast network by single cast packet of IPv4 encapsulation.

The working process of PIM-DM can be summarized as: Neighbor Discovery, Flooding-Prune, and Graft.

1. Neigh hour Discovery

When PIM-DM router is started at beginning, Hello message is required to discover neighbors. The network nodes running PIM-DM use Hello message to contact each other. PIM-DM Hello message is sent periodically.

2. Flooding-Prune

PIM-DM assumes that all hosts on the network are ready to receive multicast data. When certain multicast source S begins to send data to a multicast group G, after receiving the multicast packet, the router will make RPF examination first according to the unicast table. If the check passes, the router will create a (S, G) table item and forward the multicast packet to all downstream PIM-DM nodes (Flooding). If the RPF examination fails, i.e. the multicast packet is inputted from the incorrect interface, and then the message is discarded. After this procedure, every node will create an (S, G) item in the PIM-DM multicast domain. If there is no multicast group member in the downstream nodes, then a Prune message is sent to upstream nodes notifying not to forward data to this multicast group any more. After receiving Prune message, the corresponding interfaces will be deleted from the output interface list corresponding with the multicast-forwarding item (S, G). Through this process, a SPT (Shortest Path Tree) is established with source S as root. Prune process is started by a sub-router.

The process above is called Flooding-Prune process. Each pruned node also provides overtime mechanism at the same time. In case of overtime of prune, the router will restart flooding-prune process. Flooding-prune of PIM-DM is conducted periodically

3. RPF examination

Adopting RPF examination, PIM-DM establishes a multicast forwarding tree initiating from data source, using existing unicast routing table. When a multicast packet arrives, the router will determine the correctness of its coming path first. If the arrival interface is the interface connected to multicast source indicated by unicast routing, then this multicast packet is considered to be from the correct path; otherwise the multicast packet will be discarded as redundant message. The unicast routing message used as path judgment can root in any Unicast Routing Protocol, such as messages found by RIP, OSPF, etc. It doesn't rely on any specific unicast routing protocol.

4. Assert Mechanism

If two multicast router A and B in the same LAN segment have their own receiving paths to multicast source S, they will respectively forward multicast data packet to LAN after receiving the packet from multicast source S. Then downstream nodes multicast router C will receive two multicast packets that are exactly the same. Once router detects such circumstance, a unique forwarder will be selected through "assert" mechanism. The optimized forwarding path is selected through "assert" packet. If the priority and costs of two or more than two paths are same, the node with a larger IP address will be selected as the upstream neighbor of item (S, G), which will be responsible for forwarding the (S, G) multicast packet.

5. Graft

When the pruned downstream node needs to recover to forwarding status, this node uses Graft Message to notify upstream nodes to resume multicast data forwarding.

29.1.2 PIM-DM6 Configuration Task List

Enable PIM-DM (Required)
Configure static multicast routing entries (Optional)
Configure additional PIM-DM parameters (Optional)
Configure parameters for PIM-DM interfaces
Configure the interval for PIM-DM hello messages
Configure the interval for PIM-DM state-refresh messages
Configure the boundary interfaces
Configure the management boundary
Disable PIM-DM protocol

1. Enable the PIM-DM protocol

On QTECH series switches, PIM-DM can be enabled through two steps. Firstly PIM multicast routing should be enabled in global configuration mode, then PIM-DM should be configured for the specific interfaces.

Command	Explanation
Command configuration mode	
ipv6 pim multicast-routing	To enable PIM-DM multicast routing global. However, in order to enable PIM-DM for specific interfaces, the following command must be issued.

Enable PIM-SM for the specific interface:

Command	Explanation
Interface configuration mode	
ipv6 pim dense-mode	To enable PIM-DM for the specified interface (required).

2. Configure static multicast routing entries

Command	Explanation
Global configuration mode	
ipv6 mroute <X:X::X:X> <X:X::X:X> <ifname> <.ifname> **no ipv6 mroute <*X:X::X:X*> <*X:X::X:X*> [<*ifname*> <*.ifname*>]**	To configure IPv6 static multicast routing entries. The no form of this command will remove the specified routing entry.

3. Configure additional PIM-DM parameters

(1) Configure parameters for PIM-DM interfaces

1) Configure the interval for PIM-DM hello messages

Command	Explanation
Interface Configuration Mode	
ipv6 pim hello-interval <*interval*> **no ipv6 pim hello-interval**	To configure the interval for PIM-DM hello messages. The no form of this command will restore the default value.

2) Configure the interval for PIM-DM state-refresh messages

Command	Explanation
Interface Configuration Mode	
ipv6 pim state-refresh origination-interval **no ipv6 pim state-refresh origination-interval**	To configure the interval for sending PIM-DM state-refresh packets. The no form of this command will restore the default value.

3) Configure the boundary interfaces

Command	Explanation
Interface Configuration Mode	
ipv6 pim bsr-border **no ipv6 pim bsr-border**	To configure the interface as the boundary of PIM-DM6 protocol. On the boundary interface, STATE REFRESH messages will not be sent or received. The network connected the interface is considered as directly connected network. The no form of this command will remove the configuration.

4) Configure the management boundary

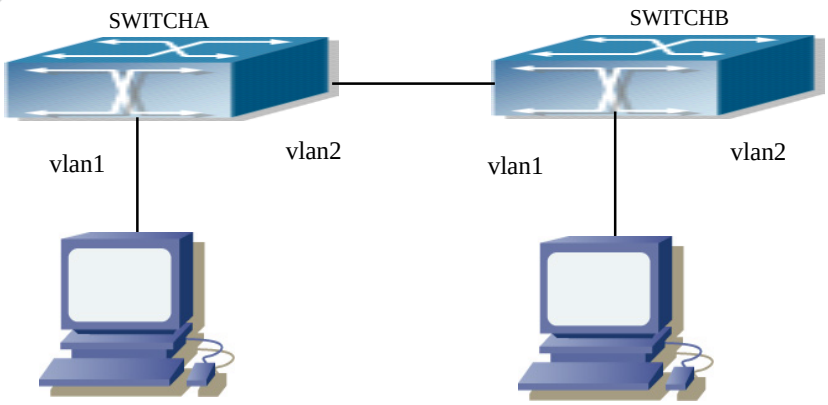
Command	Explanation
Interface Configuration Mode	
ipv6 pim scope-border *<500-599>*|*<acl_name>* **no ipv6 pim scope-border**	To configure PIM-DM6 management boundary for the interface and apply ACL for the management boundary. With default settings, ffx0::/13 is considered as the scope of the management group. If ACL is configured, then the scope specified by ACL permit command is the scope of the management group. acl_name should be standard IPv6 ACL name. The no form of this command will remove the configuration.

4. Disable PIM-DM protocol

Command	Notes
Interface Configuration Mode	
no ipv6 pim dense-mode	To disable PIM-DM for the specified interface.
Global Configuration Mode	
no ipv6 pim multicast-routing	To disable PIM-DM globally.

29.1.3 PIM-DM6 Typical Application

As shown in the following figure, add the Ethernet interfaces of Switch A and Switch B to corresponding vlan, and start PIM-DM Protocol on each vlan interface.



PIM-DM Typical Environment

The configuration procedure for SwitchA and SwitchB is as below:
(1) Configure SwitchA:
Switch(config)#ipv6 pim multicast-routing
Switch(config)#interface vlan 1
Switch(Config-if-Vlan1)#ipv6 address 2000:10:1:1::1/64
Switch(Config-if-Vlan1)#ipv6 pim dense-mode
Switch(Config-if-Vlan1)#exit
Switch(config)#interface vlan2
Switch(Config-if-Vlan2)#ipv6 address 2000:12:1:1:: 1/64
Switch(Config-if-Vlan2)#ipv6 pim dense-mode
(2) Configure SwitchB:
Switch(config)#ip pim multicast-routing
Switch(config)#interface vlan 1
Switch(Config-if-Vlan1)#ipv6 address 2000:12:1:1::2/64
Switch(Config-if-Vlan1)#ipv6 pim dense-mode
Switch(Config-if-Vlan1)#exit
Switch(config)#interface vlan 2
Switch(Config-if-Vlan2)#ipv6 address 2000:20:1:1::1/64
Switch(Config-if-Vlan2)#ipv6 pim dense-mode

29.1.4 PIM-DM6 Troubleshooting

When configuring and using PIM-DM protocol, PIM-DM protocol may fail to work normally due to physical connections, incorrect configuration and so on. So, users shall note the following points:
Assure the physical connection is correct.
Assure the Protocol of Interface and Link is UP (use show interface command);
Assure PIM Protocol is turned on in Global Mode (use ipv6 pim multicast-routing command)
Start PIM-DM Protocol on the interface (use ipv6 pim dense-mode command)
Unicast route shall be used to carry out RPF examination for multicast protocol. So the

correctness of unicast route shall be guaranteed above all. If all attempts fail to solve the problems on PIM-DM, then use debug commands such as debug ipv6 pim, copy DEBUG information in 3 minutes and send to Technology Service Center.

29.2 PIM-SM6

29.2.1 Introduction to PIM-SM6

PIM-SM6(Protocol Independent Multicast, Sparse Mode) is the IPv6 version of Protocol Independent Multicast Sparse Mode. It is a multicast routing protocol in sparse mode and mainly used in large network with group members distributed relatively sparse and wide. It is no difference from the IPv4 version PIM-SM except the addresses it uses are IPv6 addresses. Thus we don't differentiate between PIM-SM and PIM-SM6 in this chapter. All PIM-SM in the text without specific explanation is IPv6 version PIM-SM. Unlike the Flooding-Prune of Dense Mode, PIM-SM Protocol assumes no host needs receiving multicast data packets. PIM-SM router forwards multicast data packets to a host only on definite request.

By setting RP (Rendezvous Point) and BSR (Bootstrap Router), PIM-SM announce multicast packet to all PIM-SM routers and establish, using Join/Prune message of routers, RPT (RP-rooted shared tree) based on RP. Consequently the network bandwidth occupied by data packets and control messages is cut down and the transaction cost of routers is reduced. Multicast data get to the network segment where the multicast group members are located along the shared tree flow. When the data traffic reaches a certain amount, multicast data stream can be switched to source-based SPT (Shortest Path Tree) to shorten network delay. PIM-SM doesn't rely on any specific unicast routing protocol but make RPF examination using existing unicast routing table.

1. PIM-SM Working Principle

The working process of PIM-SM mainly includes neighbor discovery, creation of RPT, registration of multicast source, SPT switch and so on. The neighbor discovery mechanism is the same with the mechanism of PIM-DM. We won't introduce any more.

(1) Creation of RP Shared Tree (RPT)

When a host joins a multicast group G, the leaf router directly connected with the host finds out through IGMP message that there is a receiver of multicast group G, then it works out the corresponding Rendezvous Point RP for multicast group G, and send join message to upper level nodes in RP direction. Every router on the way from the leaf router to RP will create a (*, G) table item, indicating the message from any source to multicast group G is suitable for this item. When RP receives the message sent to multicast group G, the message will get to the leaf router along the established path and then reach the host. In this way, the RPT with RP as root is created.

(2) Multicast Source Registration

When multicast source S sends a multicast packet to multicast group G, the PIM-SM multicast router directly connected to it will take charge of sealing the multicast packet into registered message and unicast it to corresponding RP. If there are more than one PIM-SM multicast routers on a network segment, then DR (Designated Router) takes charge of forwarding the multicast packet.

(3) SPT Switch

Once the multicast router finds that the rate of the multicast packet from RP with destination address G exceeds threshold, the multicast router will send Join message to the upper lever nodes in the source direction, which results in the switch from RPT to SPT.

2. Preparation before PIM-SM configuration

(1) Configuration Candidate RP

More than one RPs (candidate RP) are permitted in PIM-SM network and each C-RP (Candidate RP) takes charge of forwarding multicast packets with destination address in a certain range. To configure more than one candidate RPs can achieve RP load balancing. There is no master or slave difference among RPs. All multicast routers work out the RP corresponded with certain multicast group based on the same algorithm after receiving the candidate RP message announced by BSR.

Note that one RP can serve more than one multicast groups, even all multicast groups. But each multicast group can only correspond with one unique RP at any moment. It can't correspond with more RPs at the same time.

(2) BSR Configuration

As the management core of PIMSM network, BSR is in charge of collecting messages sent by candidate RPs and broadcast them..

There may be only one BSR within a network. However, there may be several candidate BSRs to be configured. With such arrangement, once a BSR fails, another may be switched to. C-BSR determines BSR through automatic selection.

29.2.2 PIM-SM6 Configuration Task List

Enable PIM-SM (Required)

Configure static multicast routing entries (Optional)

Configure additional parameters for PIM-SM (Optional)

Configure parameters for PIM-SM interfaces

Configure the interval for PIM-SM hello messages

Configure the holdtime for PIM-SM hello messages

Configure ACL for PIM-SM6 neighbors

Configure the interface as the boundary interface of the PIM-SM6 protocol

Configure the interface as the management boundary of the PIM-SM6 protocol

Configure global PIM-SM parameters

Configure the switch as a candidate BSR
Configure the switch as a candidate RP
Configure static RP
Configure the cache time of kernel multicast route
Disable the PIM-SM protocol

1. Enable PIM-SM protocol

The PIM-SM protocol can be enabled on QTECH series Layer 3 switches by enabling PIM6 in global configuration mode and then enabling PIM-SM for specific interfaces in the interface configuration mode.

Command	Explanation
Global Configuration Mode	
[no] ipv6 pim multicast-routing	To enable the PIM-SM6 protocol for all the interfaces (However, in order to make PIM-SM work for specific interfaces, the following command should be issued). (required)

Make the PIM-SM protocol work for specific interfaces

Command	Explanation
Interface Configuration Mode	
[no] ipv6 pim sparse-mode [passive]	To enable PIM-SM for the specified interface. The no form of this command will disable the PIM-SM protocol (required).

2. Configure static multicast routing entries

Command	Explanation
Global Configuration Mode	
ipv6 mroute <X:X::X:X> <X:X::X:X> <ifname> <.ifname> **no ipv6 mroute <*X:X::X:X*> <*X:X::X:X*> [<*ifname*> <*.ifname*>]**	To configure a static multicast routing entry. The no form of this command will remove the specified static multicast routing entry.

3. Configure the additional parameters for PIM-SM

(1) Configure parameters for PIM-SM interfaces

1) Configure the interval for PIM-SM hello messages

Command	Explanation
Interface Configuration Mode	
ipv6 pim hello-interval <*interval*> **no ipv6 pim hello-interval**	To configure the interval for PIM-SM hello messages. The no form of this command restores

	the interval to the default value.

2) Configure the hold time for PIM-SM6 hello messages

Command	Explanation
Interface Configuration Mode	
ipv6 pim hello-holdtime <*value*> **no ipv6 pim hello-holdtime**	To configure the value of the holdtime field in the PIM-SM hello messages. The no form of this command will restore the hold time to the default value.

3) Configure ACL for PIM-SM6 neighbors

Command	Explanation
Interface Configuration Mode	
ipv6 pim neighbor-filter <*acess-list-name*> **no ipv6 pim neighbor-filter <*acess-list-name*>**	To configure ACL to filter PIM-SM6 neighbor. If session to the neighbor has been denied by ACL, then the sessions that have been set up will be discarded immediately and new sessions will not be set up.

4) Configure the interface as the boundary interface of the PIM-SM6 protocol

Command	Explanation
Interface Configuration Mode	
ipv6 pim bsr-border **no ipv6 pim bsr-border**	To configure the interface as the boundary of PIM-SM6 protocol. On the boundary interface, BSR messages will not be sent or received. The network connected the interface is considered as directly connected network. The no form of this command will remove the configuration.

5) Configure the interface as the management boundary of the PIM-SM6 protocol

Command	Explanation
Interface Configuration Mode	
ipv6 pim scope-border <*500-599*>|<*acl_name*> **no ipv6 pim scope-border**	To configure PIM-SM6 management boundary for the interface and apply ACL for the management boundary. With default settings, ffx0::/13 is considered as the scope of the management group. If ACL is configured, then the scope specified by ACL permit command is the scope of the management group. acl_name should be standard IPv6 ACL name. The no form of this command will remove the configuration.

(2) Configure global PIM-SM6 parameter

1) Configure the switch as a candidate BSR

Command	Explanation
Global Configuration Mode	
ipv6 pim bsr-candiate {vlan <*vlan_id*>|<*ifname*> | tunnel <*1-50*>}[hash-mask-length] [priority] **no Ipv6 pim bsr-candiate {vlan <*vlan_id*>|<*ifname*> | tunnel <*1-50*>}[hash-mask-length] [priority]**	This command is the global candidate BSR configuration command, which is used to configure the information of PIM-SM candidate BSR so that it can compete for BSR router with other candidate BSR. The no operation is to cancel the configuration of BSR.

2) Configure the switch as a candidate RP

Command	Explanation
Global Configuration Mode	
ipv6 pim rp-candiate {vlan<*vlan-id*>| loopback<*index*>|<*ifname*>} [<*group range*>] [<*priority*>] **no ipv6 pim rp-candiate**	This command is the global candidate RP configuration command, which is used to configure the information of PIM-SM candidate RP so that it can compete for RP router with other candidate RP. The no operation is to cancel the configuration of RP.

3) Configure static RP

Command	Explanation
Global Configuration Mode	
ipv6 pim rp-address <*rp-address*> [<*group-range*>] **no ipv6 pim rp-address <*rp-address*> {all|<*group-range*>}**	To configure the address of the candidate RP. The no form of this command will remove the configuration for the candidate RP.

4) Configure the cache time of kernel multicast route

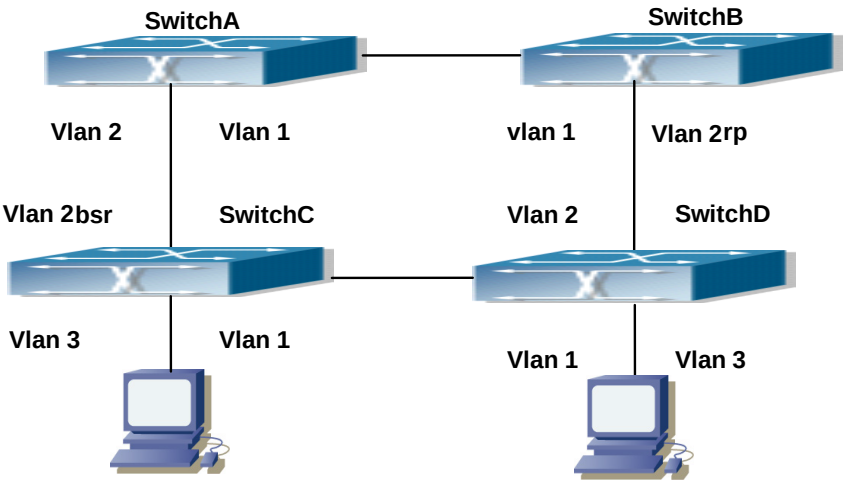
Command	Explanation
Global Configuration Mode	
ipv6 multicast unresolved-cache aging-time <value> **no ipv6 multicast unresolved-cache aging-time**	Configure the cache time of kernel multicast route, the no command restores the default value.

4. Disable PIM-SM protocol

Command	Explanation
Interface Configuration Mode	
no ipv6 pim sparse-mode	To disable the PIM-SM6 protocol.
Global Configuration Mode	
no ipv6 pim sparse-mode	To disable PIM-DM globally.

29.2.3 PIM-SM6 Typical Application

As shown in the following figure, add the Ethernet interfaces of SwitchA, SwitchB, SwitchC and SwitchD to corresponding VLAN, and start PIM-SM Protocol on each VLAN interface.



PIM-SM Typical Environment

The configuration procedure for SwitchA, SwitchB, SwitchC and SwitchD is as below:
Configure SwitchA:
Switch(config)#ipv6 pim multicast-routing
Switch(config)#interface vlan 1
Switch(Config-if-Vlan1)#ipv6 address 2000:12:1:1::1/64
Switch(Config-if-Vlan1)#ipv6 pim sparse-mode
Switch(Config-if-Vlan1)#exit
Switch(config)#interface vlan 2
Switch(Config-if-Vlan2)#ipv6 address 2000:13:1:1::1/64
Switch(Config-if-Vlan2)#ipv6 pim sparse-mode
Configure Switch B:
Switch(config)#ipv6 pim multicast-routing
Switch(config)#interface vlan 1
Switch(Config-if-Vlan1)#ipv6 address 2000:12:1:1::2/64
Switch(Config-if-Vlan1)#ipv6 pim sparse-mode
Switch(Config-if-Vlan1)#exit
Switch(config)#interface vlan 2

Switch(Config-if-Vlan2)#ipv6 address2000:24:1:1::2/64
Switch(Config-if-Vlan2)#ipv6 pim sparse-mode
Switch(Config-if-Vlan2)#exit
Switch(config)#ipv6 pim rp-candidate vlan2
Configure SwitchC:
Switch(config)#ipv6 pim multicast-routing
Switch(config)#interface vlan 1
Switch(Config-if-Vlan1)#ipv6 address 2000:34:1:1::3/64
Switch(Config-if-Vlan1)#ipv6 pim sparse-mode
Switch(Config-if-Vlan1)#exit
Switch(config)#interface vlan 2
Switch(Config-if-Vlan2)#ipv6 address 2000:13:1:1::3/64
Switch(Config-if-Vlan2)#ipv6 pim sparse-mode
Switch(Config-if-Vlan2)#exit
Switch(config)#interface vlan 3
Switch(Config-if-Vlan3)#ipv6 address 2000:30:1:1::1/64
Switch(Config-if-Vlan3)#ipv6 pim sparse-mode
Switch(Config-if-Vlan3)#exit
Switch(config)#ipv6 pim bsr-candidate vlan2 30 10
Configure SwitchD:
Switch(config)#ipv6 pim multicast-routing
Switch(config)#interface vlan 1
Switch(Config-if-Vlan1)#ipv6 address 2000:34:1:1::4/64
Switch(Config-if-Vlan1)#ipv6 pim sparse-mode
Switch(Config-if-Vlan1)#exit
Switch(config)#interface vlan 2
Switch(Config-if-Vlan2)#ipv6 address 2000:24:1:1::4/64
Switch(Config-if-Vlan2)#ipv6 pim sparse-mode
Switch(Config-if-Vlan2)#exit
Switch(config)#interface vlan 3
Switch(Config-if-Vlan3)#ipv6 address 2000:40:1:1::1/64
Switch(Config-if-Vlan3)#ipv6 pim sparse-mode

29.2.4 PIM-SM6 Troubleshooting

When configuring and using PIM-SM protocol, PIM-SM protocol may fail to work normally due to physical connections, incorrect configuration and so on. So, users shall note the following points:

Assure the physical connection is correct.

Assure the Protocol of Interface and Link is UP (use show interface command);

Unicast route shall be used to carry out RPF examination for multicast protocol. So the correctness of unicast route shall be guaranteed above all.
PIM-SM Protocol requires supports of RP and BSR, therefore you should use show ipv6 pim bsr-router first to see if there is BSR information. If not, you need to check if there is unicast routing leading to BSR.
Use show ipv6 pim rp-hash command to check if RP information is correct; if there is no RP information, you still need to check unicast routing;
If all attempts fail to solve the problems on PIM-SM, then use debug commands such as debug ipv6 pim/ debug ipv6 pim bsr, copy DEBUG information in 3 minutes and send to Technology Service Center.

29.3 ANYCAST RP v6 Configuration

29.3.1 Introduction to ANYCAST RP v6

Anycast RP v6 is a technology based on PIM protocol, which provides redundancy in order to recover as soon as possible once an RP becomes unusable.
The kernel concept of Anycast RP v6 is that the RP addresses configured all over the whole network exist on multiple multicast servers (the most common situation is that every device providing ANYCAST RP uses LOOPBACK interface, and using the longest mask to configures RP addresses on this interface), while the unicast routing algorithm will make sure that PIM routers can always find the nearest RP, thus , providing a shorter and faster way to find RP in a larger network., Once an RP being used becomes unusable, the unicast routing algorithm will ensure that the PIM router can find a new RP path fast enough to recover the multicast server in time. Multiple RP will cause a new problem that is if the multicast source and the receivers are registered to different RP, some receivers will not be able to receive data of multicast source (obviously, the register messages only prefer the nearest RP). So, in order to keep the communication between all RP, Anycast RP defines that the nearest RP to the multicast source should forward the source register messages to all the other RP to guarantee that all joiners of the RP can find the multicast source.
The method to realize the PIM-protocol-based Anycast RP is that: maintaining an ANYCAST RP list on every switch configured with Anycast RP and using another address as the label to identify each other. When one Anycast RP device receives a register message, it will send the register message to other Anycast RP devices while using its own address as the source address, to notify all the other devices of the original destination.

29.3.2 ANYCAST RP v6 Configuration Task

1. Enable ANYCAST RP v6 function
2. Configure ANYCAST RP v6

1. Enable ANYCAST RP v6 function

Command	Explanation
Global Configuration Mode	
ipv6 pim anycast-rp **no ipv6 pim anycast-rp**	Enable ANYCAST RP function. (necessary) The no operation will globally disable the ANYCAST RP function.

2. Configure ANYCAST RP v6

(1) Configure RP candidate

Command	Explanation
Global Configuration Mode	
ipv6 pim rp-candidate {vlan<*vlan-id*> |loopback<*index*> |<*ifname*>} [<*A:B::C:D*>][<*priority*>] **no ipv6 pim rp-candidate**	Now, the PIM-SM has allowed the Loopback interface to be a RP candidate.(necessary) Please pay attention to that, ANYCAST RP protocol can configure the Loopback interface or a regular three-layer VLAN interface to be the RP candidate. In make sure that PIM routers in the network can find where the RP locates, the RP candidate interface should be added into the router. No operation will cancel the RP candidate configured on this router.

(2) Configure self-rp-address (the RP communication address of this router)

Command	Explanation
Global Configuration Mode	
ipv6 pim anycast-rp self-rp-address A:B::C:D **no ipv6 pim anycast-rp self-rp-address**	Configure the self-rp-address of this router (as a RP). This address can be used to exclusively identify this router when communicating with other RP.(necessary) the effect of **self-rp-address** refers to two respects: 1 Once this router (as a RP) receives the register message from a DR unicast, it needs to forward the register message to all the other RP in the network, notifying them of the state of source (S.G). While forwarding the register message, this router will change the source address of it into self-rp-address. 2 Once this router(as a RP) receives a register message from other RP unicast, such as a register message whose destination is the self-rp-address of this router, it

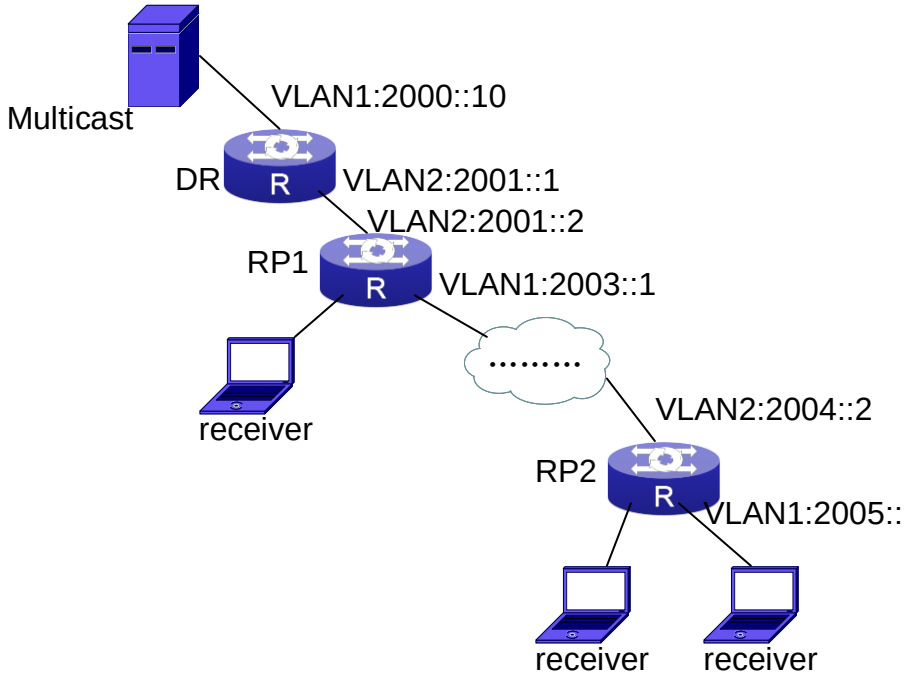
	will create (S,G) state and send back a register-terminating message, whose destination address is the source address of the register message. Pay attention: self-rp-address has to be the address of a three-layer interface on this router, but the configuration is allowed to be done with the absence of the interface. The self-rp-address should be unique. No operation will cancel the self-rp-address which is used to communicate with other RP by this router.

(3) Configure other-rp-address (other RP communication addresses)

Command	Explanation
Global Configuration Mode	
ipv6 pim anycast-rp <anycast-rp-addr> <other-rp-addr> **no ipv6 pim anycast-rp <anycast-rp-addr> <other-rp-addr>**	Configure anycast-rp-addr on this router (as a RP). This unicast address is actually the RP address configured on multiple RP in the network, in accordance with the address of RP candidate interface (or Loopback interface). The effect of **anycast-rp-addr** includes: 1 Although more than one anycast-rp-addr addresses are allowed to be configured, only the one having the same address with the currently configured RP candidate address will take effect. Only after that, can the other-rp-address in accordance with this anycast-rp-addr take effect. 2 The configuration is allowed to be done with the absence of the interface in accordance with the anycast-rp-addr. Configure on this router (as a RP) the other-rp-addresses of other RP communicating with it. This unicast address identifies other RP and is used in the communication with local routers. The effect of **other-rp-address** refers to two respects: 1 Once this router (as a RP) receives the register message from a DR unicast, it should forward it to other RP in the network to notify all the RP in the network of the source (S.G) state. While forwarding, the router will change the destination address of the register message into other-rp-address.

	2 Multiple other-rp-addresses can be configured in accordance with one anycast-rp-addr, Once the register message from a DR is received, it should be forwarded to all of this RP one by one. No operation will cancel other-rp-address communicating with this router.

29.3.3 ANYCAST RP v6 Configuration Examples



The ANYCAST RP v6 function of a router

The following is the configuration steps:

RP1 Configuration:

Switch#config

Switch(config)#interface loopback 1

Switch(Config-if-Loopback1)#ipv6 address 2006::1/128

Switch(Config-if-Loopback1)#exit

Switch(config)#ipv6 pim rp-candidate loopback1

Switch(config)#ipv6 pim bsr-candidate vlan 1

Switch(config)#ipv6 pim multicast-routing

Switch(config)#ipv6 pim anycast-rp

Switch(config)#ipv6 pim anycast-rp self-rp-address 2003::1

Switch(config)#ipv6 pim anycast-rp 2006::1 2004::2

RP2 Configuration:

Switch#config
Switch(config)#interface loopback 1
Switch(Config-if-Loopback1)#ipv6 address 2006::1/128
Switch(Config-if-Loopback1)#exit
Switch(config)#ipv6 pim rp-candidate loopback1
Switch(config)#ipv6 pim multicast-routing
Switch(config)#ipv6 pim anycast-rp
Switch(config)#ipv6 pim anycast-rp self-rp-address 2004::2
Switch(config)#ipv6 pim anycast-rp 2006::1 2003::1
Please pay attention to that, for promulgating loopback interface router, if use MBGP4+ protocol, then can use network command; or use RIPng protocol, then can use route command.

29.3.4 ANYCAST RP v6 Troubleshooting

When configuring and using ANYCAST RP v6 function, the ANYCAST RP might work abnormally because of faults in physical connections, configurations or something others. So, the users should pay attention to the following points:
The physical connections should be guaranteed to be correct
The PIM-SM6 protocol should be guaranteed to operate normally
The ANYCAST RP should be guaranteed to be enabled in Global configuration mode
The self-rp-address should be guaranteed to be configured correctly in Global configuration mode
The other-rp-address should be guaranteed to be configured correctly in Global configuration mode
All the interface routers should be guaranteed to be correctly added, including the loopback interface as a RP
Use “**show ipv6 pim anycast rp status**” command to check whether the configuration information of ANYCAST RP is correct
If the problems of ANYCAST still cannot be solved after checking, please use debug commands like “debug ipv6 pim anycast-rp”, then copy the DEBUG information within three minutes and send it to the technical service center of our company.

29.4 PIM-SSM6

29.4.1 Introduction to PIM-SSM6

Source Specific Multicast (PIM-SSM6) is a new kind of multicast service protocol. With PIM-SSM6, a multicast session is distinguished by the multicast group address and multicast source address. In SSM6, hosts can be added into the multicast group manually and efficiently

like the traditional PIM-SM6, but leave out the shared tree and RP management in PIM-S6M. In SSM6, SPT tree will be constructed with (S,G). G for the multicast group address and S for the source address of the multicast which sends datagram to G. (S,G) in a pair is named as a channel of SSM6. SSM6 serves best for the application of multicast service which is from one station to many ones, for example, the network sports video channel, and the news channel. By default, the multicast group address of SSM6 is limited to ff3x::/32. However this address range can be extended according to actual situations.

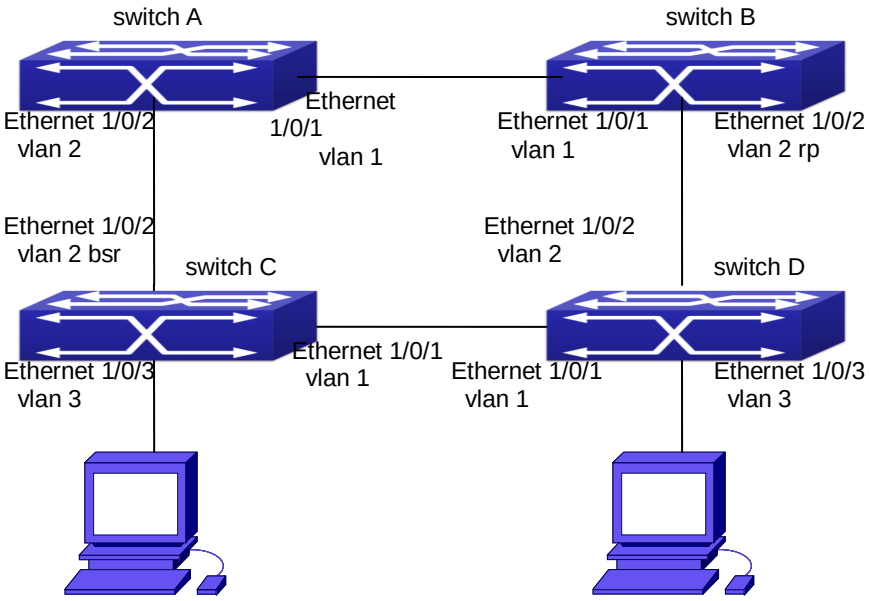
PIM-SSM6 can be supported in the PIM-DM6 environment.

29.4.2 PIM-SSM6 Configuration Task List

Command	Explanation
Global configuration mode	
ipv6 pim ssm {default|range <*access-list-number*>} **no ipv6 pim ssm**	To configure address range for pim-ssm multicast group. The no prefix will disable this command.

29.4.3 PIM-SSM6 Configuration Example

As it is shown in the below figure, ethernet interfaces of switchA, switchB, switchC, and switchD are separated into different vlan. And PIM-SM6 or PIM-DM6 is enabled on all the vlan interfaces. Take configuration of PIM-SM6 for example.



PIM-SSM typical environment

Configurations of switchA , switchB, switchC and switchD are listed as below:

(1) Configuration of switchA:

Switch(config)#ipv6 pim multicast-routing
Switch(config)#interface vlan 1
Switch(Config-If-Vlan1)# ipv6 address 2000:12:1:1::1/64
Switch(Config-If-Vlan1)# ipv6 pim sparse-mode
Switch(Config-If-Vlan1)#exit
Switch(config)#interface vlan 2
Switch(Config-If-Vlan2)# ipv6 address 2000:13:1:1::1/64
Switch(Config-If-Vlan2)# ipv6 pim sparse-mode
Switch(Config-If-Vlan2)#exit
Switch(config)#ipv6 access-list 500 permit ff1e::1/64
Switch(config)#ip pim ssm range 500

(2)Configuration of switchB:
Switch(config)#ipv6 pim multicast-routing
Switch(config)#interface vlan 1
Switch(Config-If-Vlan1)# ipv6 address 2000:12:1:1::2/64
Switch(Config-If-Vlan1)# ipv6 pim sparse-mode
Switch(Config-If-Vlan1)#exit
Switch(config)#interface vlan 2
Switch(Config-If-Vlan2)# ipv6 address2000:24:1:1::2/64
Switch(Config-If-Vlan2)# ipv6 pim sparse-mode
Switch(Config-If-Vlan2)# exit
Switch(config)# ipv6 pim rp-candidate vlan2
Switch(config)#ipv6 access-list 500 permit ff1e::1/64
Switch(config)#ip pim ssm range 500

(3) Configuration of SwitchC:
Switch(config)#ipv6 pim multicast-routing
Switch(config)#interface vlan 1
Switch(Config-If-Vlan1)# ipv6 address 2000:34:1:1::3/64
Switch(Config-If-Vlan1)# ipv6 pim sparse-mode
Switch(Config-If-Vlan1)#exit
Switch(config)#interface vlan 2
Switch(Config-If-Vlan2)# ipv6 address 2000:13:1:1::3/64
Switch(Config-If-Vlan2)# ipv6 pim sparse-mode
Switch(Config-If-Vlan2)#exit
Switch(config)#interface vlan 3
Switch(Config-If-Vlan3)# ipv6 address 2000:30:1:1::1/64
Switch(Config-If-Vlan3)# ipv6 pim sparse-mode

```
Switch(Config-If-Vlan3)# exit
Switch(config)# ipv6 pim bsr-candidate vlan2 30 10
Switch(config)#ipv6 access-list 500 permit ff1e::1/64
Switch(config)#ip pim ssm range 500
```

(4) Configuration of SwitchD:

```
Switch(config)#ipv6 pim multicast-routing
Switch(config)#interface vlan 1
Switch(Config-If-Vlan1)# ipv6 address 2000:34:1:1::4/64
Switch(Config-If-Vlan1)# ipv6 pim sparse-mode
Switch(Config-If-Vlan1)#exit
Switch(config)#interface vlan 2
Switch(Config-If-Vlan2)# ipv6 address 2000:24:1:1::4/64
Switch(Config-If-Vlan2)# ipv6 pim sparse-mode
Switch(Config-If-Vlan2)#exit
Switch(config)#interface vlan 3
Switch(Config-If-Vlan3)# ipv6 address 2000:40:1:1::1/64
Switch(Config-If-Vlan3)# ipv6 pim sparse-mode
Switch(Config-If-Vlan3)#exit
Switch(config)#ipv6 access-list 500 permit ff1e::1/64
Switch(config)#ip pim ssm range 500
```

29.4.4 PIM-SSM6 Troubleshooting

When configuring the PIM-SSM6 protocol, it may fail to work because of the failure of physical connection or the mis-configurations. To debug these errors, attention should be paid to the following lists.

Make sure the physical links are connected correctly.

Make sure the state of the data link layer has become UP. (Use show interface command).

Make sure PIM6 is enabled in global configuration mode (Refer to the command ipv6 pim multicast-routing).

Make sure PIM-SM6 is configured on the interface (Refer to the command ipv6 pim sparse-mode)

Make sure SSM6 is configure in global configuration mode.

The multicast protocol uses the unicast routing to make RPF check. Hence, single-cast routing should be verified firstly.

If problems could not be fixed with the above check list, please enable the command of **debug ipv6 pim event and debug ipv6 pim packet**, and save the debug information for 3 minutes, and send if to Technology Service Center.

29.5 IPv6 DCSCM

29.5.1 Introduction to IPv6 DCSCM

The technology of IPv6 DCSCM (Destination Control and Source Control Multicast) includes three aspects: the multicast source control, the multicast user control and the service-priority-oriented policy multicast.

IPv6 DCSCM Controllable Multicast technology proceeds as the following way:

1. If source controlled multicast is configured on the edge switches, only the multicast data of the specified group from the specified source can pass.
2. The RP switches which are the core of PIM-SM will directly send REGISTER_STOP as response to the REGISTER messages not from the specified source and specified group, and no entry is allowed to be created. (This task is implemented in the PIM-SM module).

The control of multicast users of IPv6 DCSCM technology is implemented on the basis of controlling the MLD message sent from the users, so the control module is MLD snooping and the MLD module, the control logic of which includes the following three methods: controlling according to the VLAN+MAC sending the message, controlling according to the IP address sending the message, and controlling according to the input port of the message. MLD snooping can adopts all the three methods at the same time, while the MLD module, at the third layer, can only control the IP address sending the message.

The service-priority-oriented policy multicast of IPv6 DCSCM technology adopts the following method: for the confined multicast data, the user-specified priority will be set at the access point, enabling the data can be sent at a higher priority through TRUNK, and guaranteeing that the data can be sent through the whole net at the user-specified priority.

29.5.2 IPv6 DCSCM Configuration Task Sequence

The source control configuration

The destination control configuration

The multicast policy configuration

The source control configuration

The source control configuration has three steps, first is globally enabling the source control, the following is the command of globally enabling the source control:

Command	Explanation
Global Configuration Mode	
ipv6 multicast source-control(necessary) **no ipv6 multicast source-control**	Globally enable the source control, the no operation of this command will globally disable the source control. What should be paid attention to is that, once globally enable the source control, all the multicast messages will

	be dropped by default. All the source control configurations can only be done after globally enabled, and only when all the configured rules are disabled, the source control can be disabled globally.

The next is configuring the source control rules, which adopts the same method as configuring ACL, using ACL number from 8000 to 8099, while each rule number can configure 10 rules. What should be paid attention to is that these rules have orders, the earliest configured rule is at the front. Once a rule is matched, the following ones will not take effect, so the globally enabled rules should be the last to configure. The following is the command:

Command	Explanation
Global Configuration Mode	
[no] ipv6 access-list *<8000-8099>* {deny|permit} {{*<source/M>*}|{host-source *<source-host-ip>*}|any-source} {{*<destination/M>* }|{host-destination *<destination-host-ip>*}|any-destination}	Used to configure the source control rules, the rules can only take effect when applied to the specified port. The no operation of this command can delete the specified rule.

The last is to configure the rules to the specified port.

Pay attention: since the configured rules will take up entries of hardware, configuring too many rules might cause failure if the underlying entries are full, so it is recommended that users adopt rules as simple as possible. The following is the configuration command:

Command	Explanation
Port Configuration Mode	
[no] ipv6 multicast source-control access-group *<8000-8099>*	Used to configure the source control rule to a port, he no operation will cancel this configuration.

The configuration of destination control

The configuration of destination control is similar to that of source control, and also has three steps:

First, globally enable the destination control, since destination control needs to avoid the unauthorized users from receiving multicast data, once it is enabled globally, the switch will stop broadcasting received multicast data, so if a switch has enabled destination control, users should not connect two or more other Layer three switches within the same VLAN where it locates. The following is the configuration command:

Command	Explanation
Global Configuration Mode	
multicast destination-	Globally enable IPV4 and IPv6 destination control,

control(necessary)	the no operation of this command will globally disable destination control. All of the other configuration can only take effect after globally enabled.

The next is configuring destination control rules, which are similar to that of source control, but using ACL number from 9000 to 10099 instead.

Command	Explanation
Global Configuration Mode	
[no] ipv6 access-list <*9000-10099*> {deny|permit} {{<*source/M*>}|{host-source <*source-host-ip*>}|any-source} {{<*destination/M*>}|{host-destination <*destination-host-ip*>}|any-destination}	Used to configure destination control rules, these rules can only take effect when applied to specified source IP, VLAN-MAC or port. The no operation of this rule will delete the specified rule.

The last step is to configure the rules to the specified source IP, source VLAN MAC or the specified port. What should be paid attention to is that only when the MLD-SNOOPING is enabled, these rules can be globally used, or, only rules of source IP can be used in MLD protocol. The following is the configuration command:

Command	Explanation
Port Mode	
[no] ipv6 multicast destination-control access-group <*9000-10099*>	Used to configure the destination control rule to a port, the no operation of this command will cancel the configuration.
Global Configuration Mode	
[no] ipv6 multicast destination-control <*1-4094*> <*macaddr*> access-group <*9000-10099*>	Used to configure the destination control rules to the specified VLAN-MAC, the no operation of this command will cancel the configuration.
[no] ipv6 multicast destination-control <*IPADDRESS/M*> access-group <*9000-100999*>	Used to configure the destination control rules to the specified source IPv6 address/MASK, the no operation of this command will cancel the configuration.

The configuration of multicast policy

The multicast policy adopts the method of specifying a priority for the specified multicast data to meet the user's particular demand, what should be paid attention to is that only when multicast data is transmitted in TRUNK, can it be taken special care of. The configuration is quite simple, for only one command is needed, that is set priority for the specified multicast, the following is the command:

Command	Explanation
Global Configuration Mode	
[no] ipv6 multicast policy <IPADDRESS/M> <IPADDRESS/M> cos <priority>	Configure multicast policy, set priority for sources and groups in a specified range, the priority valid range is 0 to 7.

29.5.3 IPv6 DCSCM Typical Examples

1. Source control

In order to prevent an edge switch sends multicast data at will, we configure on the edge switch that only the switch whose port is Ethernet1/0/5 can send multicast data, and the group of data should be ff1e::1. The uplink port Ethernet1/0/25 can forward multicast data without being restricted, so we can configure as follows.

Switch(config)#ipv6 access-list 8000 permit any-source ff1e::1
Switch(config)#ipv6 access-list 8001 permit any any
Switch(config)#ipv6 multicast source-control
Switch(config)#interface Ethernet1/0/5
Switch(Config-If-Ethernet1/0/5)#ipv6 multicast source-control access-group 8000
Switch(config)#interface Ethernet1/0/25
Switch(Config-If-Ethernet1/0/25)#ipv6 multicast source-control access-group 8001

2. Destination control

We want to confine that the users of the segment whose address is fe80::203:fff:fe01:228a/64 can not join the ff1e::1/64 group, so we can configure as follows:

First, enable MLD Snooping in the VLAN where it locates (in this example, it is VLAN2).

Switch(config)#ipv6 mld snooping
Switch(config)#ipv6 mld snooping vlan 2

Then configure relative destination control access list and configure specified IPv6 address to use this access list.

Switch(config)#ipv6 access-list 9000 deny any ff1e::1/64
Switch(config)#ipv6 access-list 9000 permit any any
Switch(config)#multicast destination-control
Switch(config)#ipv6 multicast destination-control fe80::203:fff:fe01:228a/64 access-group 9000

Thus, the users of this segment can only join groups other than 2ff1e::1/64.

3. Multicast policy

Server 2008::1 is sending important multicast data in group ff1e::1, we can configure on its access switch as follows:

Switch(config)#ipv6 multicast policy 2008::1/128 ff1e::1/128 cos 4

Thus this multicast flow will have a priority of 4, when it passes the TRUNK port of this switch to another switch (generally speaking, it is a relatively high priority, the data with higher priority might be protocol data, if a higher priority is set, when there is too much multicast data, the

switch protocol might operate abnormally).

29.5.4 IPv6 DCSCM Troubleshooting

IPv6 DCSCM module acts like ACL, so most problems are caused by improper configuration. Please read the instructions above carefully.

29.6 MLD

29.6.1 Introduction to MLD

MLD (Multicast Listener Discovery) is the multicast group member (receiver) discovery protocol serving IPv6 multicast. It is similar to IGMP Protocol in IPv4 multicast application. Correspondingly, MLD Protocol version1 is similar to IGMP Protocol version2, and MLD Protocol version2 is similar to IGMP Protocol version3. Current firmware supports MLDv1/ MLDv2.

The IPv6 multicast hosts can join or leave from multicast group at any location, any time, regardless of the total number of group members. It is unnecessary and impossible for multicast switch to store the relationship among all host members. Multicast switch simply finds out via MLD protocol if there are receivers of certain multicast group on the network segment connected to each port. The only thing host need to do is to keep the record of which multicast groups it joined.

MLD is unsymmetrical between host and switch: the host needs to respond the MLD query message of multicast switch with membership report message; the switch periodically sends membership query message and determines if there is host joining a specific group in its subnetworks according to the response message received, and after it receives the report of a host quitting from the group, it sends out the query for the group to confirm if there is no member left in it.

There are three types of protocol messages of MLD Protocol, that is, Query, Report and Done (which is corresponding to Leave of IGMPv2). Like IGMPV2, the Query messages include General Query and Specific Group Query. General Query uses the multicast address FF02::1 of hosts as destination address, the group address is 0; and Specific Group Query use its group address as destination address. The multicast addresses of MLD use 130, 131 and 132 as data types denoting the three kinds of messages mentioned above. Other logic is basically same as IGMPv2.

MLD protocol version2 use FF02::16 as destination address of membership report, and 143 as data type. The other logic of MLD Protocol version2 is similar to IGMP Protocol version3.

29.6.2 MLD Configuration Task List

Start MLD (Required)

Configure MLD auxiliary parameters (Required)

(1) Configure MLD group parameters

1) Configure MLD group filter conditions

(2) Configure MLD query parameters

1) Configure the interval of MLD sending query message

2) Configure the maximum response time of MLD query

3) Configure overtime of MLD query

Shut down MLD Protocol

Start MLD Protocol

There is no special command for starting MLD Protocol on EDGECORE series layer 3 switches. MLD Protocol will automatically start up as long as any IPv6 multicast protocol is started on corresponding interface.

Command	Explanation
Global Mode	
ipv6 pim multicast-routing	To start Global IPv6 Multicast Protocol, the precondition of starting MLD Protocol. The NO operation of corresponding command shuts ipv6 multicast protocol and MLD Protocol. (Required)

Command	Explanation
Port Configuration Mode	
ipv6 pim dense-mode | ipv6 pim sparse-mode	Start MLD Protocol. The NO operation of corresponding command shuts MLD Protocol. (Required)

Configure MLD auxiliary parameters

(1) Configure MLD group parameters

Configure MLD group filter conditions

Command	Explanation
Port Configuration Mode	
ipv6 mld access-group *<acl_name>* no ipv6 mld access-group	Configure the filter conditions of interface for MLD group; the NO operation of this command cancels filter conditions.

(2) Configure MLD Query parameters

1) Configure interval time for MLD to send query messages

2) Configure the maximum response time of MLD query

3) Configure the overtime of MLD query

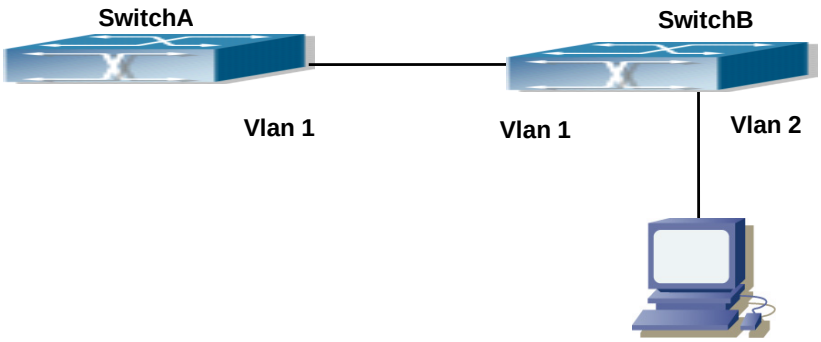
Command	Explanation
Port Configuration Mode	
ipv6 mld query-interval <*time_val*> **no ipv6 mld query-interval**	Configure the interval of MLD query messages sent periodically; the NO operation of this command restores the default value.
ipv6 mld query-max-response-time <*time_val*> **no ipv6 mld query-max-response-time**	Configure the maximum response time of the interface for MLD query; the NO operation of this command restores the default value.
ipv6 mld query-timeout <*time_val*> **no ipv6 mld query-timeout**	Configure the overtime of the interface for MLD query; the NO operation of this command restores the default value.

Shut down MLD Protocol

Command	Explanation
Port Configuration Mode	
no ipv6 pim dense-mode | no ipv6 pim sparse-mode | no ipv6 pim multicast-routing (Global Mode)	Shut down MLD Protocol

29.6.3 MLD Typical Application

As shown in the following figure, add the Ethernet interfaces of Switch A and Switch B to corresponding vlan, and start PIM6 on each vlan interface.



Network Topology Diagram

The configuration procedure for SwitchA and SwitchB is as below:

(1) Configure SwitchA:

Switch (config) #ipv6 pim multicast-routing

Switch (config) #ipv6 pim rp-address 3FFE::1

Switch (config) #interface vlan 1

Switch (Config-if-Vlan1) #ipv6 address 3FFE::1/64

Switch (Config-if-Vlan1) #ipv6 pim sparse-mode

(2) Configure SwitchB:

Switch (config) #ipv6 pim multicast-routing
Switch (config) #ipv6 pim rp-address 3FFE::1
Switch (config) #interface vlan1
Switch (Config-if-Vlan1) #ipv6 address 3FFE::2/64
Switch (Config-if-Vlan1) #ipv6 pim sparse-mode
Switch (Config-if-Vlan1) #exit
Switch (config) #interface vlan2
Switch (Config-if-Vlan2) #ipv6 address 3FFA::1/64
Switch (Config-if-Vlan2) #ipv6 pim sparse-mode
Switch (Config-if-Vlan2) #ipv6 mld query-timeout 150

29.6.4 MLD Troubleshooting Help

When configuring and using MLD protocol, MLD protocol may fail to work normally due to physical connections, incorrect configuration and so on. So, users shall note the following points:

Assure the physical connection is correct.

Assure the protocol of interface and link is UP (use show interface command)

Assure to start one kind of multicast protocol on the interface

Assure the time of the timers of each router on the same network segment is consistent; usually we recommend the default setting.

Unicast route shall be used to carry out RPF examination for multicast protocol. So the correctness of unicast route shall be guaranteed above all.

If all attempts fail to solve the problems on MLD, please use debug commands such as debug ipv6 MLD event/packet, and copy DEBUG information in 3 minutes and send to Technology Service Center.

29.7 MLD Snooping

29.7.1 Introduction to MLD Snooping

MLD, the Multicast Listener Discovery Protocol, is used to realize multicasting in the IPv6. MLD is used by the network equipments such as routers which supports multicast for multicast listener discovery, also used by listeners looking forward to join certain multicast group informing the router to receive data packets from certain multicast address, all of which are done through MLD message exchange. First the router send an MLD Multicast listener Query message through a multicast address which can address all the listeners (namely ff02::1). Once there is a listener who wishes to join the multicast address, it will send a MLD Multicast listener Report back through the multicast address.

MLD Snooping is namely the MLD listening. The switch restricts the multicast traffic from flooding through MLD Snooping, and forward the multicast traffic to ports associated to

multicast devices only. The switch listens to the MLD messages between multicast routers and listeners, and maintains the multicast group forwarding list based on the listening result. The switches forwards multicast packets according to the multicast forwarding list

The switch realizes the MLD Snooping function while supporting MLD v2. This way, the user can acquire IPv6 multicast with the switch.

29.7.2 MLD Snooping Configuration Task

1. Enable the MLD Snooping function
2. Configure the MLD Snooping

1. Enable the MLD Snooping function

Command	Explanation
Global Mode	
ipv6 mld snooping **no ipv6 mld snooping**	Enable global MLD Snooping, the "**no ipv6 mld snooping**" command disables the global MLD snooping.

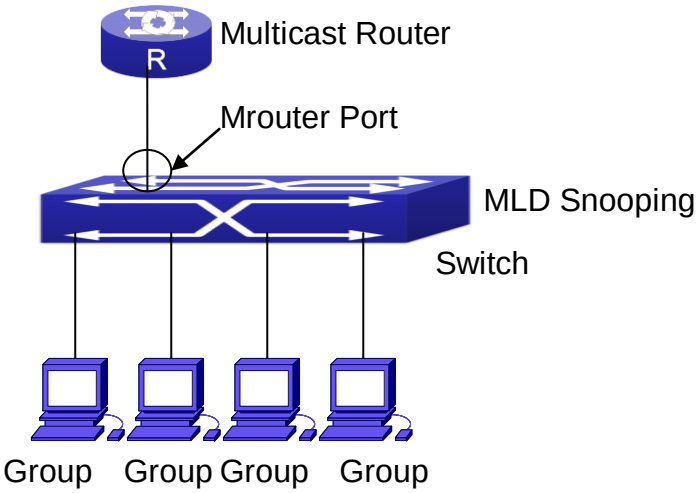
2. Configure MLD Snooping

Command	Explanation
Global Mode	
ipv6 mld snooping vlan <*vlan-id*> **no ipv6 mld snooping vlan <*vlan-id*>**	Enable MLD Snooping on specific VLAN. The "no" form of this command disables MLD Snooping on specific VLAN.
ipv6 mld snooping vlan <*vlan-id*> limit {group <*g_limit*> | source <*s_limit*>} **no ipv6 mld snooping vlan <*vlan-id*> limit**	Configure the number of the groups in which the MLD Snooping can join, and the maximum number of sources in each group. The "no" form of this command restores to the default.
ipv6 mld snooping vlan <*vlan-id*> l2-general-querier **no ipv6 mld snooping vlan <*vlan-id*> l2-general-querier**	Set the VLAN level 2 general querier, which is recommended on each segment. The "no" form of this command cancels the level 2 general querier configuration.
ipv6 mld snooping vlan <*vlan-id*> mrouter-port interface <*interface –name*> **no ipv6 mld snooping vlan <*vlan-id*> mrouter-port interface <*interface –name*>**	Configure the static mrouter port in specific vlan. The "no" form of this command cancels the mrouter port configuration.
ipv6 mld snooping vlan <*vlan-id*> mrouter-port learnpim6 **no ipv6 mld snooping vlan <*vlan-id*>**	Enable the function that the specified VLAN learns mrouter-port (according to pimv6 packets), the no command will disable the

mrouter-port learnpim6	function.
ipv6 mld snooping vlan <*vlan-id*> mrpt <*value*> **no ipv6 mld snooping vlan <*vlan-id*> mrpt**	Configure the keep-alive time of the mrouter port. The “no” form of this command restores to the default.
ipv6 mld snooping vlan <*vlan-id*> query-interval <*value*> **no ipv6 mld snooping vlan <*vlan-id*> query-interval**	Configure the query interval. The “no” form of this command restores to the default.
ipv6 mld snooping vlan <*vlan-id*> immediate-leave **no ipv6 mld snooping vlan <*vlan-id*> immediate-leave**	Configure immediate leave multicast group function for the MLD Snooping of specific VLAN. The “no” form of this command cancels the immediate leave configuration.
ipv6 mld snooping vlan <*vlan-id*> query-mrsp <*value*> **no ipv6 mld snooping vlan <*vlan-id*> query-mrsp**	Configure the query maximum response period. The “no” form of this command restores to the default.
ipv6 mld snooping vlan <*vlan-id*> query-robustness <*value*> **no ipv6 mld snooping vlan <*vlan-id*> query-robustness**	Configure the query robustness, the “no” form of this command restores to the default.
ipv6 mld snooping vlan <*vlan-id*> suppression-query-time <*value*> **no ipv6 mld snooping vlan <*vlan-id*> suppression-query-time**	Configure the suppression query time. The “no” form of this command restores to the default
ipv6 mld snooping vlan <vlan-id> static-group <X:X::X:X> [source <X:X::X:X>] interface [ethernet | port-channel] <IFNAME> **no ipv6 mld snooping vlan <vlan-id> static-group <X:X::X:X> [source <X:X::X:X>] interface [ethernet | port-channel] <IFNAME>**	Configure static-group on specified port of the VLAN. The no form of the command cancels this configuration.

29.7.3 MLD Snooping Examples

Scenario 1: MLD Snooping Function



Open the switch MLD Snooping Function figure

As shown above, the vlan 100 configured on the switch consists of ports 1, 2, 6, 10, 12. Four hosts are respectively connected to 2, 6, 10, 12 while the multicast router on port 1. Suppose we need MLD Snooping on VLAN 100, however by default, the global MLD Snooping as well as the MLD Snooping on each VLAN are, therefore first we have to enable the global MLD Snooping at the same time enable the MLD Snooping on VLAN 100, furthermore we need to set the port 1 of VLAN 100 as a mrouter port.

Configuration procedure is as follows.

Switch#config

Switch(config)#ipv6 mld snooping

Switch(config)#ipv6 mld snooping vlan 100

Switch(config)#ipv6 mld snooping vlan 100 mrouter-port interface ethernet 1/0/1

Multicast configuration:

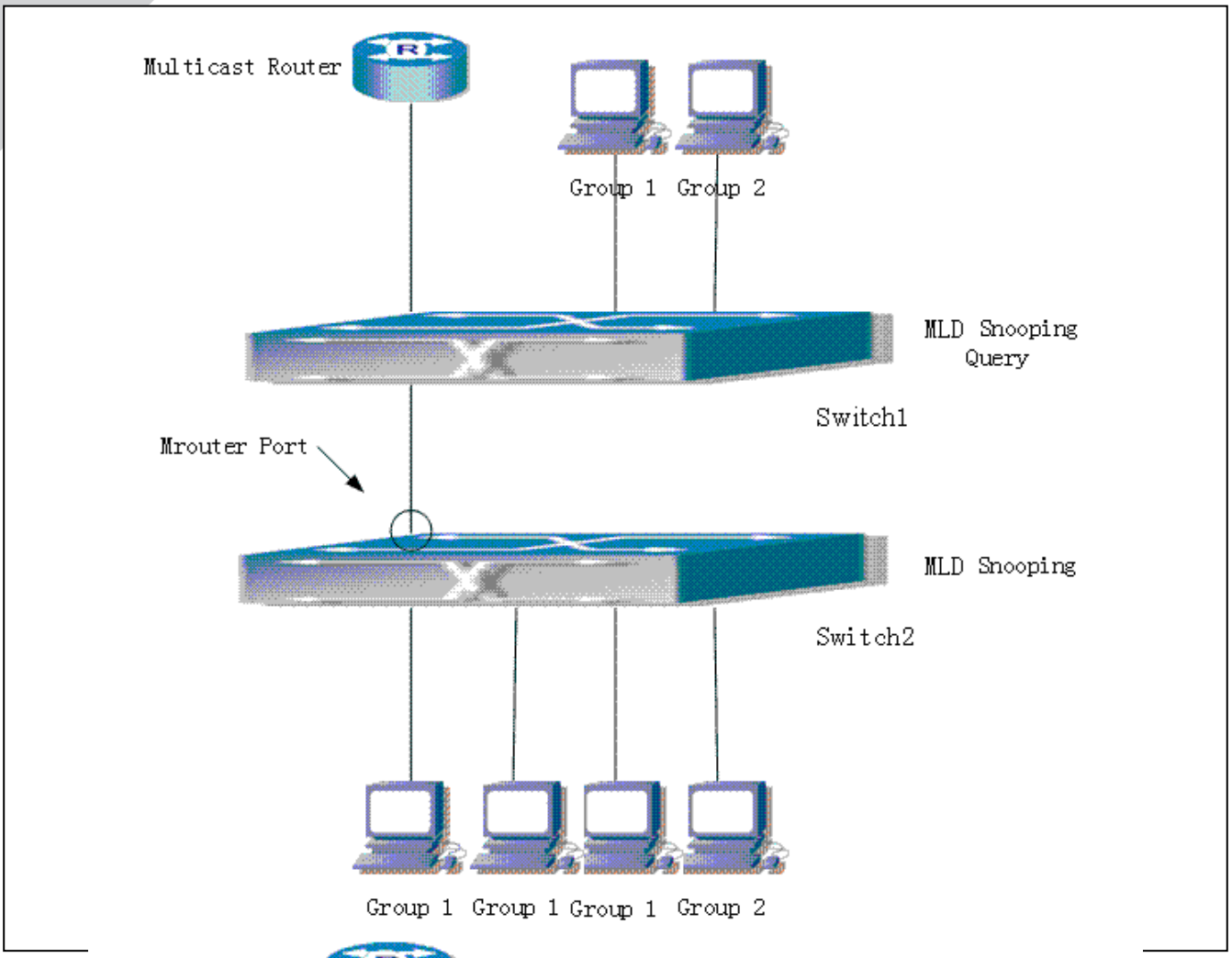
Assume there are two multicast servers: the Multicast Server 1 and the Multicast Server 2, amongst program 1 and 2 are supplied on the Multicast Server 1 while program 3 on the Multicast server 2, using group addresses respectively the Group 1, Group 2 and Group 3. Concurrently multicast application is operating on the four hosts. Two hosts connected to port 2 and 5 are playing program 1 while the host connected to port 10 playing program 2, and the one to port 12 playing program 3.

MLD Snooping interception results:

The multicast table on vlan 100 shows: port1, 2 and 6 are in (Multicasting Server 1, Group1) , port1, 10 are in (Multicasting Server 1,Group2), and port1, 12 are in (Multicasting Server 2, Group3)

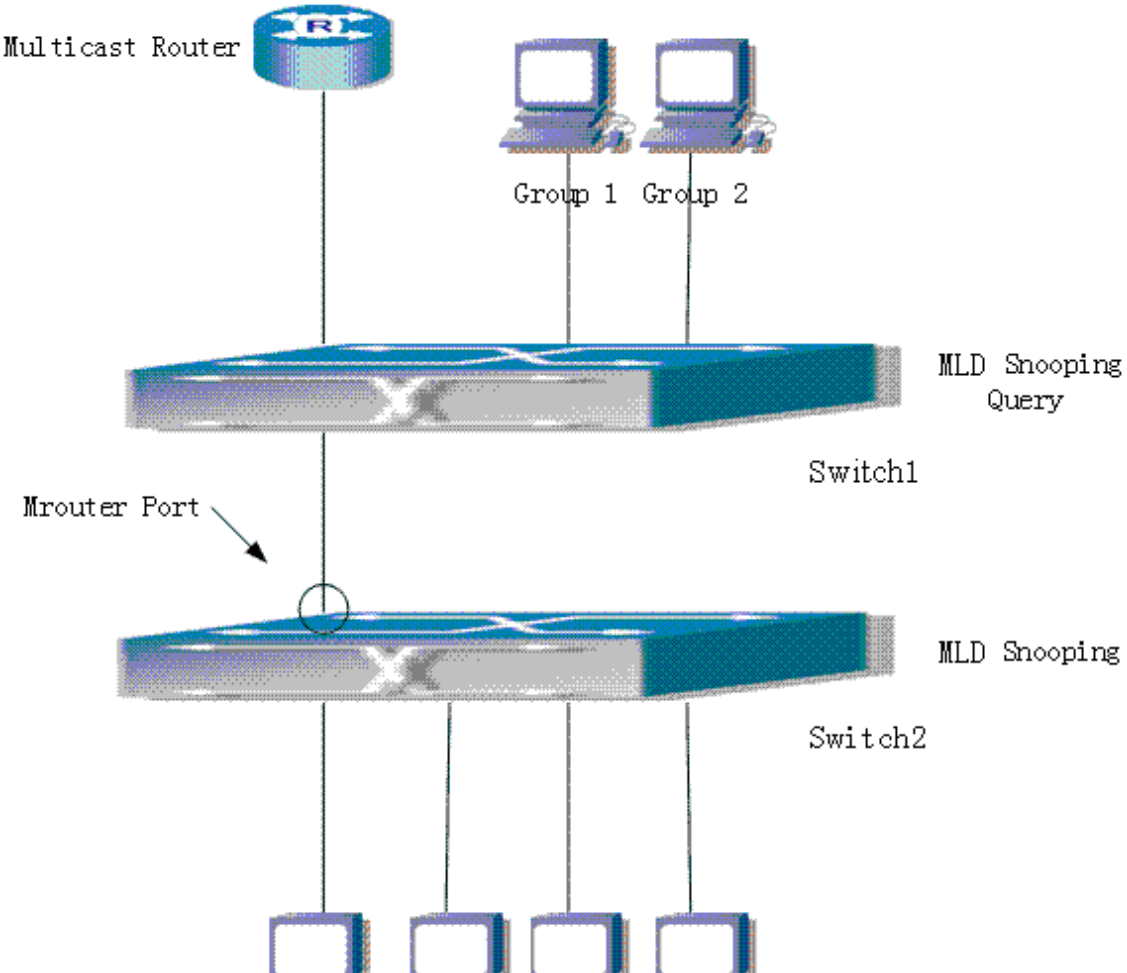
All the four hosts successfully receive programs they are interested in. port2, 6 receives no traffic from program2 and 3; port10 receives no traffic from program 1 and 3, and port12 receives no traffic from program1 and 2.

Scenario 2: MLD L2-general-querier



Switch
Config
replace
10, 12
periodi
l2-gene
Config
Switch
Switch
Switch
Switch
Switch
Switch
Switch
Switch
Switch
Multica
Same

witch 1
ort 1, 2,
Query
vlan 60



MLD Snooping interception results:
Same as scenario 1

Scenario 3: To run in cooperation with layer 3 multicast protocols
switch which is used in Scenario 1 is replaced with ROUTER with specific configurations remains the same. And multicast and IGMP snooping configurations are the same with what it is in Scenario 1. To configure PIM-SM6 on ROUTER, and enable PIM-SM6 on vlan 100 (use the same PIM mode with the connected multicast router)

The configurations are listed as below:
switch#config
switch(config)#ipv6 pim multicast-routing
switch(config)#interface vlan 100
switch(config-if-vlan100)#ipv6 pim sparse-mode

MLD snooping does not distribute entries when layer 3 multicast protocol is enabled. It only does the following tasks.
To remove the layer 2 multicast entries.
To provide query functions to the layer 3 with vlan, S, and G as the parameters.
When layer 3 MLD is disabled, re-enable distributing layer 2 multicast entries.
By looking up the layer 3 IP6MC entries, it can be found that ports can be indicated by the layer 3 multicast entries. This ensures the MLD Snooping can work in cooperation with the layer 3 multicast protocols.

29.7.4 MLD Snooping Troubleshooting

In configuring and using MLD Snooping, the MLD Snooping server may fail to run properly due to physical connection failure, wrong configuration, etc. The user should ensure the following:
Ensure the physical connection is correct
Ensure the MLD Snooping is enabled under global mode (using ipv6 mld snooping)
Ensure the MLD Snooping is configured on the vlan under global mode (using ipv6 mld snooping vlan <vlan-id>)
Ensure there is a vlan configured as a L2 general querier, or there is a static mrouter configured in a segment,
Use command to check if the MLD snooping information is correct

Chapter 30 Multicast VLAN

30.1 Introductions to Multicast VLAN

Based on current multicast order method, when orders from users in different VLAN, each VLAN will copy a multicast traffic in this VLAN, which is a great waste of the bandwidth. By configuration of the multicast VLAN, we add the switch port to the multicast VLAN, with the IGMP Snooping/MLD Snooping functions enabled, users from different VLAN will share the same multicast VLAN. The multicast traffic only exists within a multicast VLAN, so the bandwidth is saved. As the multicast VLAN is absolutely separated from the user VLAN, security and bandwidth concerns can be met at the same time, after the multicast VLAN is configured, the multicast traffic will be continuously sent to the users.

30.2 Multicast VLAN Configuration Task List

Enable the multicast VLAN function
Configure the IGMP Snooping
Configure the MLD Snooping

1. Enable the multicast VLAN function

Command	Explanation
VLAN configuration mode	
multicast-vlan **no multicast-vlan**	Configure a VLAN and enable the multicast VLAN on it. The "**no multicast-vlan**" command disables the multicast function on the VLAN.
multicast-vlan association <*vlan-list*> **no multicast-vlan association <*vlan-list*>**	Associate a multicast VLAN with several VLANs. The "no" form of this command deletes the related VLANs associated with the multicast VLAN.

2. Configure the IGMP Snooping

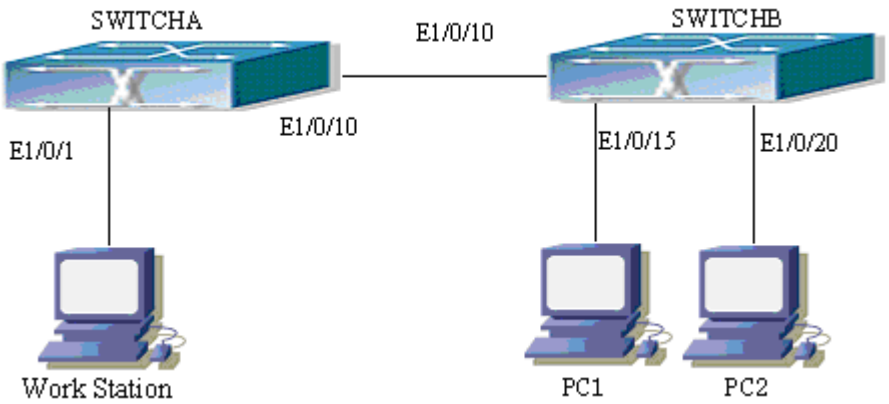
Command	Explanation
Global Mode	
ip igmp snooping vlan <*vlan-id*> **no ip igmp snooping vlan <*vlan-id*>**	Enable the IGMP Snooping function on the multicast VLAN. The "no" form of this command disables the IGMP Snooping on the multicast VLAN.

ip igmp snooping **no ip igmp snooping**	Enable the IGMP Snooping function. The “no” form of this command disables the IGMP snooping function.

3. Configure the MLD Snooping

ipv6 mld snooping vlan *<vlan-id>* **no ipv6 mld snooping vlan *<vlan-id>***	Enable MLD Snooping on multicast VLAN; the “no” form of this command disables MLD Snooping on multicast VLAN.
ipv6 mld snooping **no ipv6 mld snooping**	Enable the MLD Snooping function. The “no” form of this command disables the MLD snooping function.

30.3 Multicast VLAN Examples



Function configuration of the Multicast VLAN

As shown in the figure, the multicast server is connected to the layer 3 switch switchA through port 1/0/1 which belongs to the VLAN10 of the switch. The layer 3 switch switchA is connected with layer 2 switches through the port1/0/10, which configured as trunk port. On the switchB the VLAN100 is configured set to contain port1/0/15, and VLAN101 to contain port1/0/20. PC1 and PC2 are respectively connected to port 1/0/15 and1/0/20. The switchB is connected with the switchA through port1/0/10, which configured as trunk port. VLAN 20 is a multicast VLAN. By configuring multicast vlan, the PC1 and PC2 will receives the multicast data from the multicast VLAN.

Following configuration is based on the IP address of the switch has been configured and all the equipment are connected correctly.

Configuration procedure

SwitchA#config

SwitchA(config)#vlan 10

SwitchA(config-vlan10)#switchport access ethernet 1/0/1

SwitchA(config-vlan10)exit
SwitchA(config)#interface vlan 10
Switch(Config-if-Vlan10)#ip pim dense-mode
Switch(Config-if-Vlan10)#exit
SwitchA(config)#vlan 20
SwitchA(config-vlan20)#exit
SwitchA(config)#interface vlan 20
SwitchA(Config-if-Vlan20)#ip pim dense-mode
SwitchA(Config-if-Vlan20)#exit
SwitchA(config)#ip pim multicast
SwitchA(config)# interface ethernet1/0/10
SwitchA(Config-If-Ethernet1/0/10)switchport mode trunk

SwitchB#config
SwitchB(config)#vlan 100
SwitchB(config-vlan100)#Switchport access ethernet 1/0/15
SwitchB(config-vlan100)exit
SwitchB(config)#vlan 101
SwitchB(config-vlan101)#Switchport access ethernet 1/0/20
SwitchB(config-vlan101)exit
SwitchB(config)# interface ethernet 1/0/10
SwitchB(Config-If-Ethernet1/0/10)#Switchport mode trunk
SwitchB(Config-If-Ethernet1/0/10)#exit
SwitchB(config)#vlan 20
SwitchB(config-vlan20)#multicast-vlan
SwitchB(config-vlan20)#multicast-vlan association 100,101
SwitchB(config-vlan20)#exit
SwitchB(config)#ip igmp snooping
SwitchB(config)#ip igmp snooping vlan 20

When the multicast VLAN supports the IPv6 multicast, the usage is the same with IPv4, but the difference is using with MLD Snooping, so does not give an example.

Chapter 31 VRRP Configuration

31.1 Introduction to VRRP

VRRP (Virtual Router Redundancy Protocol) is a fault tolerant protocol designed to enhance connection reliability between routers (or L3 Ethernet switches) and external devices. It is developed by the IETF for local area networks (LAN) with multicast/broadcast capability (Ethernet is a Configuration Example) and has wide applications.

All hosts in one LAN generally have a default route configured to specified default gateway, any packet destined to an address outside the native segment will be sent to the default gateway via this default route. These hosts in the LAN can communicate with the external networks. However, if the communication link connecting the router serving as default game and external networks fails, all hosts using that gateway as the default next hop route will be unable to communicate with the external networks.

VRRP emerged to resolve such problem. VRRP runs on multiple routers in a LAN, simulating a "virtual" router (also referred to as a "Standby cluster") with the multiple routes. There is an active router (the "Master") and one or more backup routers (the "Backup") in the Standby cluster. The workload of the virtual router is actually undertaken by the active router, while the Backup routers serve as backups for the active router.

The virtual router has its own "virtual" IP address (can be identical with the IP address of some router in the Standby cluster), and routers in the Standby cluster also have their own IP address. Since VRRP runs on routes or Ethernet Switches only, the Standby cluster is transparent to the hosts with the segment. To them, there exists only the IP address of the Virtual Router instead of the actual IP addresses of the Master and Backup(s). And the default gateway setting of all the hosts uses the IP address of the Virtual Router. Therefore, hosts within the LAN communicate with the other networks via this Virtual Router. But basically, they are communicating with the other networks via the Master. In the case when the Master of the Standby cluster fails, a backup will take over its task and become the Master to serve all the hosts in the LAN, so that uninterrupted communication between LAN hosts and external networks can be achieved.

To sum it up, in a VRRP Standby cluster, there is always a router/Ethernet serving as the active router (Master), while the rest of the Standby cluster servers act as the backup router(s) (Backup, can be multiple) and monitor the activity of Master all the time. Should the Master fail, a new Master will be elected by all the Backups to take over the work and continue serving the hosts within the segment. Since the election and take-over duration is brief and smooth, hosts within the segment can use the Virtual Router as normal and uninterrupted communication can be achieved.

31.2 VRRP Configuration Task List

Configuration Task List:
Create/Remove the Virtual Router (required)
Configure VRRP dummy IP and interface (required)
Activate/Deactivate Virtual Router (required)
Configure VRRP sub-parameters (optional)
Configure the preemptive mode for VRRP
Configure VRRP priority
Configure VRRP Timer intervals
Configure VRRP interface monitor

1. Create/Remove the Virtual Router

Command	Explanation
Global Mode	
router vrrp <*vrid*> **no router vrrp <*vrid*>**	Creates/Removes the Virtual Router.

2. Configure VRRP Dummy IP Address and Interface

Command	Explanation
VRRP protocol configuration mode	
virtual-ip <*ip*> **no virtual-ip**	Configures VRRP Dummy IP address; the **"no virtual-ip"** command removes the virtual IP address.
interface {IFNAME | Vlan <*ID*>} **no interface**	Configures VRRP interface, the "**no interface**" command removes the interface.

3. Activate/Deactivate Virtual Router

Command	Explanation
VRRP protocol configuration mode	
enable	Activates the Virtual Router.
disable	Deactivates the Virtual Router.

4. Configure VRRP Sub-parameters

(1) Configure the preemptive mode for VRRP

Command	Explanation
VRRP protocol configuration mode	
preempt-mode {true| false}	Configures the preemptive mode for VRRP.

(2) Configure VRRP priority

Command	Explanation
VRRP protocol configuration mode	
priority <*priority*>	Configures VRRP priority.

(3) Configure VRRP Timer intervals

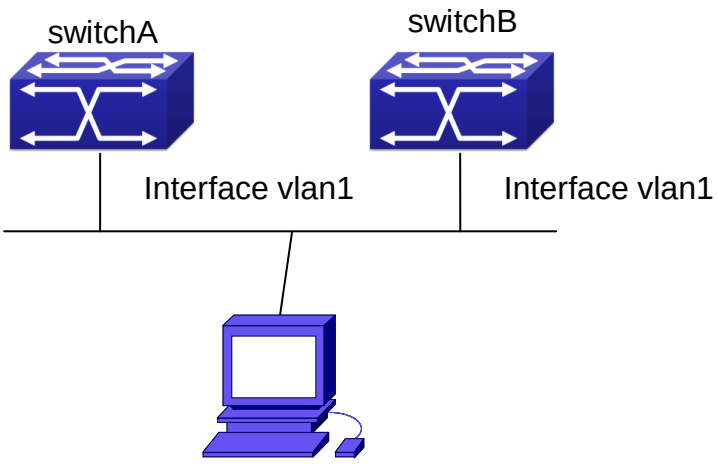
Command	Explanation
VRRP protocol configuration mode	
advertisement-interval <*time*>	Configures VRRP timer value (in seconds).

(4) Configure VRRP interface monitor

Command	Explanation
VRRP protocol configuration mode	
circuit-failover {IFNAME | Vlan <*ID*> } <*value_reduced*> **no circuit-failover**	Configures VRRP interface monitor, the **"no circuit-failover"** removes monitor to the interface.

31.3 VRRP Typical Examples

As shown in the figure below, SwitchA and SwitchB are Layer three Ethernet Switches in the same group and provide redundancy for each other.



VRRP Network Topology

Configuration of SwitchA:

SwitchA(config)#interface vlan 1

SwitchA (Config-if-Vlan1)# ip address 10.1.1.1 255.255.255.0

SwitchA (config)#router vrrp 1

SwitchA(Config-Router-Vrrp)# virtual-ip 10.1.1.5

SwitchA(Config-Router-Vrrp)# interface vlan 1

SwitchA(Config-Router-Vrrp)# enable

Configuration of SwitchB:
SwitchB(config)#interface vlan 1
SwitchB (Config-if-Vlan1)# ip address 10.1.1.7 255.255.255.0
SwitchB(config)#router vrrp 1
SwitchB (Config-Router-Vrrp)# virtual-ip 10.1.1.5
SwitchB(Config-Router-Vrrp)# interface vlan 1
SwitchB(Config-Router-Vrrp)# enable

31.4 VRRP Troubleshooting

In configuring and using VRRP protocol, the VRRP protocol may fail to run properly due to reasons such as physical connection failure or wrong configurations. The user should ensure the following:
Good condition of the physical connection.
All interface and link protocols are in the UP state (use "**show interface**" command).
Ensure VRRP is enabled on the interface. Verify the authentication mode of different routers (or L3 Ethernet switches) in the same standby cluster are the same.
Verify the timer time of different routers (or L3 Ethernet switches) in the same standby cluster are the same.
Verify the dummy IP address is in the same network segment of the interface's actual IP address.
If the examination remains unsolved, please use **debug vrrp** and other debugging command and copy the DEBUG message within 3 minutes, send the recorded message to the technical server center of our company.

Chapter 32 IPv6 VRRPv3 Configuration

32.1 Introduction to VRRPv3

VRRPv3 is a virtual router redundancy protocol for IPv6. It is designed based on VRRP (VRRPv2) in IPv4 environment. The following is a brief introduction to it.

In a network based on TCP/IP protocol, in order to guarantee the communication between the devices which are not physically connected, routers should be specified. At present there are two most commonly used methods to specify routers: one is to study dynamically via routing protocols (such as internal routing protocols RIP and OSPF); the other is to configure statically. Running dynamical routing protocol on each terminal is unrealistic, since most operating systems for client end do not support dynamical routing protocol, even if they do, they are limited by the overheads of management, convergence, security and many other problems. So the common method is to adopt static routing configuration on terminal IP devices, which usually means specify one or more default gateway for terminal devices. Static routing simplifies the management of network and reduces the communication overheads of terminal devices, but it still has a disadvantage: if the router acting as the default gateway breaks, the communication of all the hosts which use this gateway as their next hop host. Even if there are more than one default gateways, before rebooting the terminal devices, they can not switch to the new gateway. Adopting virtual router redundancy protocol (VRPR) can effectively avoid the flaws of statically specifying gateways.

In VRRP protocol, there are two groups of import concepts: VRRP routers and virtual routers, master routers and backup routers. VRRP routers are routers running VRRP, which are physical entities; virtual routers are the ones created by VRRP, which are logical concepts. A group of VRRP routers cooperate to comprise a virtual router, which acts outwardly as a logical router with a unique fixed IP address and MAC address. The routers belonging to the same VRRP group play two mutually exclusive roles at the same time: master routers and backup routers. One VRRP group can only have one master router other but one or more backup routers. VRRPv3 protocol uses selection policy to select a master router from the router group to take charge of responding ND(Neighbor Discovery) neighbor request messages(ARP in IPv4) and forwarding IP data packets, while the other routers in the group will be in a state of waiting as backups. When the master router has a problem for some season, the backup router will be updated to the master router after a delay of a few seconds. Since this switch is very fast and does not need to change IP address or MAC address, it will be transparent to terminal user systems.

In IPv6 environment, the hosts in a LAN usually learn the default gateway via neighbor discovery protocol (NDP), which is implemented based on regularly receiving advertisement messages from routers. The NDP of IPv6 has a mechanism called Neighbor Unreachability

Detection, which checks whether a neighbor node is failed by sending unicast neighbor request messages to it. In order to reduce the overheads of sending neighbor request messages, these messages are only sent to those neighbor nodes which are sending flows, and are only sent if there is no instruction of UP state of the router in a period of time. In Neighbor Unreachability Detection, if adopting default parameters, it will take about 38 seconds to detect an unreachable router, which is a delay not ignorable for users and might cause a time-out in some transport protocols. Compared with NDP, VRRP provides a fast default gateway switch. In VRRP, backup routers can take up the unavailable master router in about 3 seconds (default parameter), and this process needs no interaction with hosts, which means being transparent to hosts.

32.1.1 The Format of VRRPv3 Message

VRRPv3 has its own message format, VRRP messages are used to communicate the priority of routers and the state of Master in the backup group, they are encapsulated in IPv6 messages to send, and are sent to the specified IPv6 multicast address. The format of VRRPv3 message is shown in Graph 1. The source address of the IPv6 message encapsulating the VRRPv3 message is the local address of the outbound interface of the message, and the destination address of it is the IPv6 multicast address(the multicast allocated to VRRPv3 is FF02:0:0:0:0:0:0:12). The number of hops should be limited to 255, and the next message head is 112(representing a VRRP message).

The meaning of each field in a VRRPv3 message is shown as follows:

Version: The version of VRRPv3, whose value is 3;

Type: The type of VRRP messages. There is only one type: ADVERTISEMENT, and its value is 1;

Virtual Rtr ID: The ID of the virtual router;

Priority: Priority, ranging from 0 to 255;

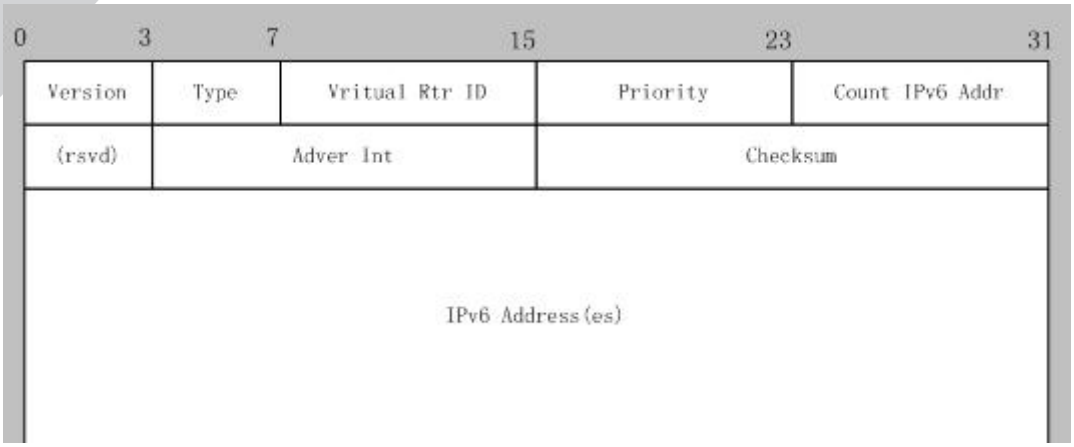
Count IPv6 Addr: The number of IPv6 addresses in a VRRPv3 message, the minimum of which is 1;

Rsvd: Reserved field, whose value is 0;

Adver Int: The advertisement interval of VRRPv3 messages, in seconds;

Checksum: The checksum, taking account of the whole VRRPv3 message and an IPv6 pseudo head (please refer to RFC2460 for details);

IPv6 Address(es): one or more IPv6 addresses related to the virtual router, the number of which is the same with ”Count IPv6 Addr”, and the first one of which should be the virtual IPv6 address of the virtual router.



VRRPv3 message

32.1.2 VRRPv3 Working Mechanism

The working mechanism of VRRPv3 is the same with that of VRRPv2, which is mainly implemented via the interaction of VRRP advertisement messages. It will be briefly described as follows:

Each VRRP router has a unique ID: VRIP, ranging from 1 to 255. This router has a unique virtual MAC address outwardly, and the format of which is 00-00-5E-00-02-{VRID} (the format of virtual MAC address in VRRPv2 is 00-00-5E-00-01-{VRID}). Master router is in charge of using this MAC address to respond to ND neighbor request (it is ARP request in VRRPv2). Thus, no matter what switch is made, the terminal devices will get the same IP and MAC address all the time, reducing the affection that the switch causes on terminal devices.

There is only one kind of VRRP control message: VRRP advertisement. It uses IP multicast data packets to encapsulate, and the format of multicast addresses is FF02:0:0:0:0:0:XXXX:XXXX. In order to keep a consistence with the multicast address in VRRPv2 (224.0.0.18), the multicast addresses used by VRRPv3 advertisement messages can be FF02:0:0:0:0:0:0:12, and the advertisement is limited within the same LAN. Thus, different VRID are guaranteed to be used repeatedly in different networks. In order to reduce the overheads of network bandwidth, only master routers can send VRRP advertisement messages regularly. Backup routers will start a new round of VRRP selection if it hasn't received a VRRP advertisement in 3 advertisement intervals in a row or if it receives an advertisement with a priority of 0.

In a VRRP router group, the master router is selected according to priority. The range of priority in VRRP protocol is 0-255. If the IP address of a VRRP router is the same to that of the virtual router interface, then the virtual router will be called the IP address owner in the VRRP group; the IP address owner automatically has the highest priority: 255. The priority of 0 is usually used when the IP address owner gives up the role of master. The range of priority can be configured is 1-254. The configuration rule of priority can be set according to the speed and cost of the link, the performance and reliability of the router and other management policies. In the selection of the master router, the virtual router with high priority will win. So, if there is an

IP owner in the VRRP group, it will always be the master router. For the candidate routers having the same priority, selection will be done according to the magnitude of IP addresses (the bigger IP address takes precedence). VRRP also provides a preemptive priority policy. If such policy is configured, the backup router with higher priority will preempt the role of new master router over the current master router with lower priority.

In order to avoid the fault of returning a physical MAC address when Pinging virtual IP, it is regulated that virtual IP can not be the real IP of the interface. Thus, all the interfaces participating of the backup group selection will be backup by default.

32.2 VRRPv3 Configuration

32.2.1 Configuration Task Sequence

Create/delete the virtual router (necessary)

Configure the virtual IPv6 address and interface of VRRPv3 (necessary)

Enable/disable the virtual router (necessary)

Configure VRRPv3 assistant parameters (optional)

(1) Configure VRRPv3 preempt mode

(2) Configure VRRPv3 priority

(3) Configure the VRRPv3 advertisement interval

(4) Configure the monitor interface of VRRPv3

1. Create/delete the virtual router

Command	Explanation
Global Configuration Mode	
router ipv6 vrrp <*vrid*> **no router ipv6 vrrp <*vrid*>**	Create/delete the virtual router.

2. Configure the virtual IPv6 address and interface of VRRPv3

Command	Explanation
VRRPv3 Protocol Mode	
virtual-ipv6 <*ipv6-address*> Interface {Vlan <*ID*> | *IFNAME* } **no virtual-ipv6 interface**	Configure the virtual IPv6 address and interface of VRRPv3, the no operation of this command will delete the virtual IPv6 address and interface.

3. Enable/disable the virtual router

Command	Explanation
VRRPv3 Protocol Mode	
enable	Enable the virtual router.

disable	Disable the virtual router.

4. Configure VRRPv3 assistant parameters

(1) Configure VRRPv3 preempt mode

Command	Explanation
VRRPv3 Protocol Mode	
preempt-mode {true| false}	Configure VRRPv3 preempt mode.

(2) Configure VRRPv3 priority

Command	Explanation
VRRPv3 Protocol Mode	
priority < *priority* >	Configure VRRPv3 priority.

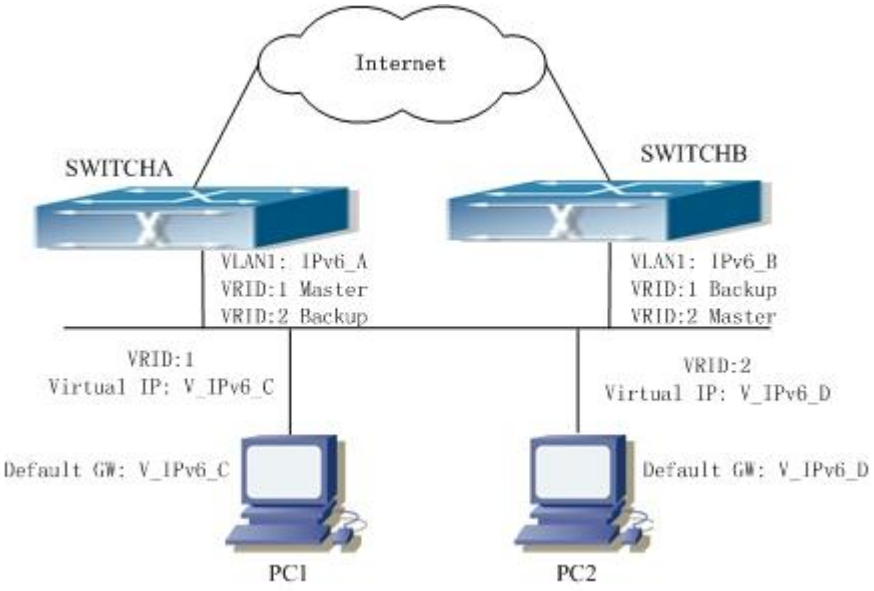
(3) Configure the VRRPv3 advertisement interval

Command	Explanation
VRRPv3 Protocol Mode	
advertisement-interval <time>	Configure the VRRPv3 advertisement interval (in cent seconds).

(4) Configure the monitor interface of VRRPv3

Command	Explanation
VRRPv3 Protocol Mode	
circuit-failover {vlan <*ID*>| *IFNAME*} <*value_reduced*> **no circuit-failover**	Configure the monitor interface of VRRPv3, the no operation of this command will delete the monitor interface.

32.3 VRRPv3 Typical Examples



VRRPv3 Typical Network Topology

As shown in graph, switch A and switch B are backups to each other, switch A is the master of backup group 1 and a backup of backup group 2. Switch B is the master of backup group 2 and a Backup of backup group 1. The IPv6 addresses of switch A and switch B are “IPv6_A” and “IPv6_ B” respectively (it is recommended that IPv6_A and IPv6_B are in the same segment), the virtual IPv6 address of backup group 1 and backup group are “V_IPv6_C” and “V_IPV6_D” respectively, and the default IPv6 gateway address are configured as “V_IPv6_C” and “V_IPv6_D” respectively (in reality, the IPv6 gateway address of hosts are usually learnt automatically via router advertisements, thus, the IPv6 next hop of the hosts will have some randomness). Doing this will not only implement router backup but also the flow sharing function in the LAN.

The configuration of SwitchA:

```
SwitchA (config)#ipv6 enable
SwitchA (config)#interface vlan 1
SwitchA (config)#router ipv6 vrrp 1
SwitchA (config-router)#virtual-ipv6 fe80::2 interface vlan 1
SwitchA (config-router)#priority 150
SwitchA (config-router)#enable
SwitchA (config)#router ipv6 vrrp 2
SwitchA (config-router)#virtual-ipv6 fe80::3 interface vlan 1
SwitchA (config-router)#enable
```

The configuration of SwitchB:

```
SwitchB (config)# ipv6 enable
SwitchB (config)# interface vlan 1
SwitchB (config)# router ipv6 vrrp 2
SwitchB (config-router)# virtual-ipv6 fe80::3 interface vlan 1
SwitchB (config-router)# priority 150
SwitchB (config-router)# enable
SwitchB (config)# router ipv6 vrrp 1
SwitchB (config-router)# virtual-ipv6 fe80::2 interface vlan 1
SwitchB (config-router)# enable
```

32.4 VRRPv3 Troubleshooting

When configuring and using VRRPv3 protocol, it might operate abnormally because of incorrect physical connections and configuration. So, users should pay attention to the following points:

First, the physical connections should be correct;

Next, the interface and link protocol are UP (use **show ipv6 interface** command);

And then, make sure that IPv6 forwarding function is enabled (use **ipv6 enable** command);
Besides, make sure that VRRPv3 protocol is enable on the interface;
Check whether the time of timer in different routers (or layer-three Ethernet switch) within the same backup group is the same;
Check whether the virtual IPv6 addresses in the same backup group is the same.

Chapter 33 MRPP Configuration

33.1 Introduction to MRPP

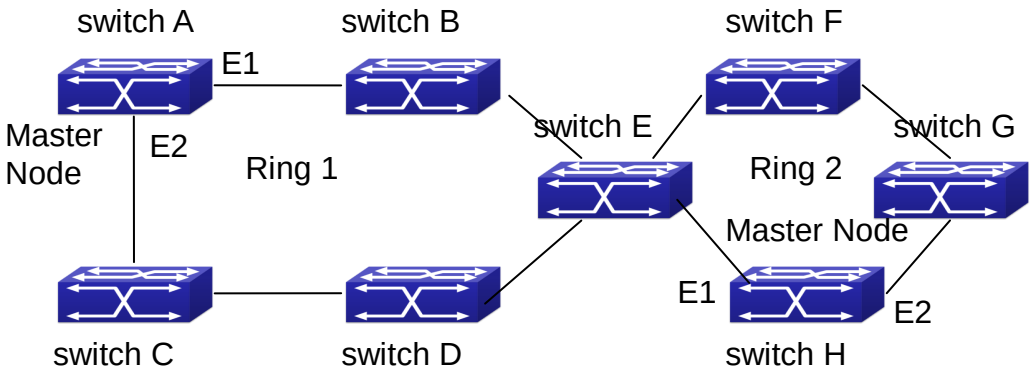
MRPP (Multi-layer Ring Protection Protocol), is a link layer protocol applied on Ethernet loop protection. It can avoid broadcast storm caused by data loop on Ethernet ring, and restore communication among every node on ring network when the Ethernet ring has a break link. MRPP is the expansion of EAPS (Ethernet link automatic protection protocol).

MRPP protocol is similar to STP protocol on function, MRPP has below characters, compare to STP protocol:

<1> MRPP specifically uses to Ethernet ring topology

<2> fast convergence, less than 1 s. ideally it can reach 100-50 ms.

33.1.1 Conception Introduction



MRPP Sketch Map

1. Control VLAN

Control VLAN is a virtual VLAN, only used to identify MRPP protocol packet transferred in the link. To avoid confusion with other configured VLAN, avoids configuring control VLAN ID to be the same with other configured VLAN ID. The different MRPP ring should configure the different control VLAN ID.

2. Ethernet Ring (MRPP Ring)

Ring linked Ethernet network topology.

Each MRPP ring has two states.

Health state: The whole ring net work physical link is connected.

Break state: one or a few physical link break in ring network

3. nodes

Each switch is named after a node on Ethernet. The node has some types:
Primary node: each ring has a primary node, it is main node to detect and defend.
Transfer node: except for primary node, other nodes are transfer nodes on each ring.
The node role is determined by user configuration. As shown Fig 3-1, Switch A is primary node of Ring 1, Switch B. Switch C; Switch D and Switch E are transfer nodes of Ring 1.

4. Primary port and secondary port
The primary node and transfer node have two ports connecting to Ethernet separately, one is primary port, and another is secondary port. The role of port is determined by user configuration.
Primary port and secondary port of primary node
The primary port of primary node is used to send ring health examine packet (hello), the secondary port is used to receive Hello packet sending from primary node. When the Ethernet is in health state, the secondary port of primary node blocks other data in logical and only MRPP packet can pass. When the Ethernet is in break state, the secondary port of primary node releases block state, and forwards data packets.
There are no difference on function between Primary port and secondary port of transfer node. The role of port is determined by user configuration. As shown Figure, Switch A E1 is primary port, E2 is secondary port.

5. Timer
The two timers are used when the primary node sends and receives MRPP protocol packet: Hello timer and Fail Timer.
Hello timer: define timer of time interval of health examine packet sending by primary node primary port.
Fail timer: define timer of overtime interval of health examine packet receiving by primary node primary port. The value of Fail timer must be more than or equal to the 3 times of value of Hello timer.

33.1.2 MRPP Protocol Packet Types

Packet Type	Explanation
Hello packet (Health examine packet) Hello	The primary port of primary node evokes to detect ring, if the secondary port of primary node can receive Hello packet in configured overtime, so the ring is normal.
LINK-DOWN (link Down event packet)	After transfer node detects Down event on port, immediately sends LINK-DOWN packet to primary node, and inform primary node ring to fail.

LINK-DOWN-FLUSH_FDB packet	After primary node detects ring failure or receives LINK-DOWN packet, open blocked secondary port, and then uses two ports to send the packet, to inform each transfer node to refresh own MAC address.
LINK-UP-FLUSH_FDB packet	After primary detects ring failure to restore normal, and uses packet from primary port, and informs each transfer node to refresh own MAC address.

33.1.3 MRPP Protocol Operation System

1. Link Down Alarm System

When transfer node finds themselves belonging to MRPP ring port Down, it sends link Down packet to primary node immediately. The primary node receives link down packet and immediately releases block state of secondary port, and sends LINK-DOWN-FLUSH-FDB packet to inform all of transfer nodes, refreshing own MAC address forward list.

2. Poll System

The primary port of primary node sends Hello packet to its neighbors timely according to configured Hello-timer.

If the ring is health, the secondary port of primary node receives health detect packet, and the primary node keeps secondary port.

If the ring is break, the secondary port of primary node can't receive health detect packet when timer is over time. The primary releases the secondary port block state, and sends LINK-DOWN-FLUSH_FDB packet to inform all of transfer nodes, to refresh own MAC address forward list.

3. Ring Restore

After the primary node occur ring fail, if the secondary port receives Hello packet sending from primary node, the ring has been restored, at the same time the primary node block its secondary port, and sends its neighbor LINK-UP-Flush-FDB packet.

After MRPP ring port refresh UP on transfer node, the primary node maybe find ring restore after a while. For the normal data VLAN, the network maybe forms a temporary ring and creates broadcast storm. To avoid temporary ring, transfer node finds it to connect to ring network port to refresh UP, immediately block temporarily (only permit control VLAN packet pass), after only receiving LINK-UP-FLUSH-FDB packet from primary node, and releases the port block state.

33.2 MRPP Configuration Task List

1) Globally enable MRPP

2) Configure MRPP ring

3) Configure the query time of MRPP

4) Display and debug MRPP relevant information

1) Globally enable MRPP

Command	Explanation
Global Mode	
mrpp enable **no mrpp enable**	Globally enable and disable MRPP.

2) Configure MRPP ring

Command	Explanation
Global Mode	
mrpp ring <*ring-id*> **no mrpp ring <*ring-id*>**	Create MRPP ring. The “no” command deletes MRPP ring and its configuration.
MRPP ring mode	
control-vlan <*vid*> **no control-vlan**	Configure control VLAN ID, format “no” deletes configured control VLAN ID.
node-mode {master | transit}	Configure node type of MRPP ring (primary node or secondary node).
hello-timer < *timer*> **no hello-timer**	Configure Hello packet timer sending from primary node of MRPP ring, format “no” restores default timer value.
fail-timer <*timer*> **no fail-timer**	Configure Hello packet overtime timer sending from primary node of MRPP ring, format “no” restores default timer value.
enable **no enable**	Enable MRPP ring, format “no” disables enabled MRPP ring.
Port mode	
mrpp ring <*ring-id*> primary-port **no mrpp ring <*ring-id*> primary-port**	Specify primary port of MRPP ring.
mrpp ring <*ring-id*> secondary-port **no mrpp ring <*ring-id*> secondary-port**	Specify secondary port of MRPP ring.

3) Configure the query time of MRPP

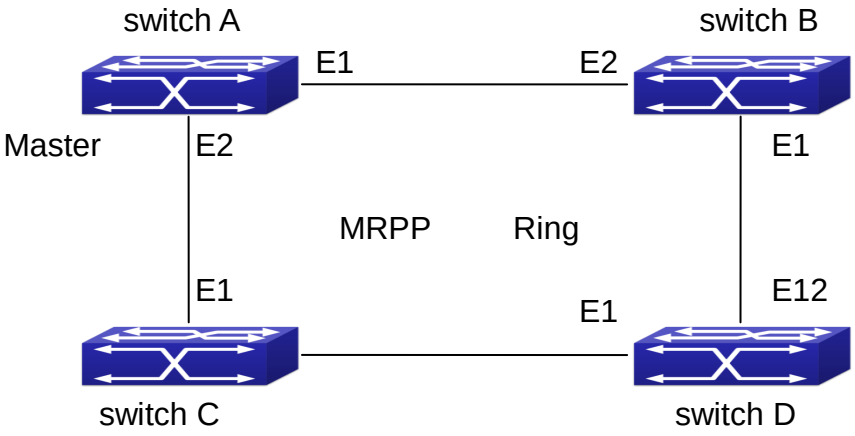
Command	Explanation
Global Mode	

mrpp poll-time <20-2000>	Configure the query interval of MRPP.

4) Display and debug MRPP relevant information

Command	Explanation
Admin Mode	
debug mrpp **no debug mrpp**	Disable MRPP module debug information, format “no” disable MRPP debug information output.
show mrpp {<*ring-id*>}	Display MRPP ring configuration information.
show mrpp statistics {<*ring-id*>}	Display receiving data packet statistic information of MRPP ring.
clear mrpp statistics {<*ring-id*>}	Clear receiving data packet statistic information of MRPP ring.

33.3 MRPP Typical Scenario



MRPP typical configuration scenario

The above topology often occurs on using MRPP protocol. The multi switch constitutes a single MRPP ring, all of the switches only are configured an MRPP ring 4000, thereby constitutes a single MRPP ring.

In above configuration, switch A configuration is primary node of MRPP ring 4000, and configures E1/0/1 to primary port, E1/0/2 to secondary port. Other switches are secondary nodes of MRPP ring, configures primary port and secondary port separately.

To avoid ring, it should temporarily disable one of the ports of primary node, when it enables each MRPP ring in the whole MRPP ring; and after all of the nodes are configured, open the port.

When disable MRPP ring, it needs to insure the MRPP ring doesn’t have ring.

switch A configuration Task Sequence:

Switch(Config)#mrpp enable

```
Switch(Config)#mrpp ring 4000
Switch(mrpp-ring-4000)#control-vlan 4000
Switch(mrpp-ring-4000)#fail-timer 18
Switch(mrpp-ring-4000)#hello-timer 5
Switch(mrpp-ring-4000)#node-mode master
Switch(mrpp-ring-4000)#enable
Switch(mrpp-ring-4000)#exit
Switch(Config)#interface ethernet 1/0/1
Switch(config-If-Ethernet1/0/1)#mrpp ring 4000 primary-port
Switch(config-If-Ethernet1/0/1)#interface ethernet 1/0/2
Switch(config-If-Ethernet1/0/2)#mrpp ring 4000 secondary-port
Switch(config-If-Ethernet1/0/2)#exit
Switch(Config)#
```

switch B configuration Task Sequence:

```
Switch(Config)#mrpp enable
Switch(Config)#mrpp ring 4000
Switch(mrpp-ring-4000)#control-vlan 4000
Switch(mrpp-ring-4000)#enable
Switch(mrpp-ring-4000)#exit
Switch(Config)#interface ethernet 1/0/1
Switch(config-If-Ethernet1/0/1)#mrpp ring 4000 primary-port
Switch(config-If-Ethernet1/0/1)#interface ethernet 1/0/2
Switch(config-If-Ethernet1/0/2)#mrpp ring 4000 secondary-port
Switch(config-If-Ethernet1/0/2)#exit
Switch(Config)#
```

switch C configuration Task Sequence:

```
Switch(Config)#mrpp enable
Switch(Config)#mrpp ring 4000
Switch(mrpp-ring-4000)#control-vlan 4000
Switch(mrpp-ring-4000)#enable
Switch(mrpp-ring-4000)#exit
Switch(Config)#interface ethernet 1/0/1
Switch(config-If-Ethernet1/0/1)#mrpp ring 4000 primary-port
Switch(config-If-Ethernet1/0/1)#interface ethernet 1/0/2
Switch(config-If-Ethernet1/0/2)#mrpp ring 4000 secondary-port
Switch(config-If-Ethernet1/0/2)#exit
Switch(Config)#
```

switch D configuration Task Sequence:
Switch(Config)#mrpp enable
Switch(Config)#mrpp ring 4000
Switch(mrpp-ring-4000)#control-vlan 4000
Switch(mrpp-ring-4000)#enable
Switch(mrpp-ring-4000)#exit
Switch(Config)#interface ethernet 1/0/1
Switch(config-If-Ethernet1/0/1)#mrpp ring 4000 primary-port
Switch(config-If-Ethernet1/0/1)#interface ethernet 1/0/2
Switch(config-If-Ethernet1/0/2)#mrpp ring 4000 secondary-port
Switch(config-If-Ethernet1/0/2)#exit
Switch(Config)#

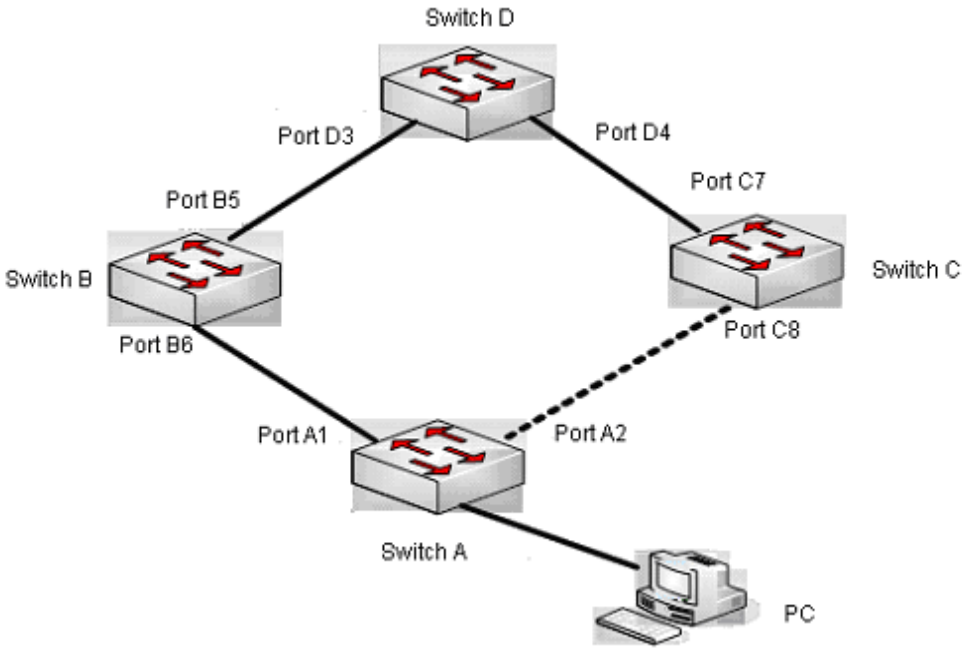
33.4 MRPP Troubleshooting

The normal operation of MRPP protocol depends on normal configuration of each switch on MRPP ring, otherwise it is very possible to form ring and broadcast storm:
Configuring MRPP ring, you'd better disconnected the ring, and wait for each switch configuration, then open the ring.
When the MRPP ring of enabled switch is disabled on MRPP ring, it ensures the ring of the MRPP ring has been disconnected.
When there is broadcast storm on MRPP ring, it disconnects the ring firstly, and ensures if each switch MRPP ring configuration on the ring is correct or not; if correct, restores the ring, and then observes the ring is normal or not.
The convergence time of MRPP ring net is relative to the response mode of up/down. If use poll mode, the convergence time as hundreds of milliseconds in simple ring net, if use interrupt mode, the convergence time within 50 milliseconds.
Generally, the port is configured as poll mode, interrupt mode is only applied to better performance environment, but the security of poll mode is better than interrupt mode, port-scan-mode {interrupt | poll} command can be consulted.
In normal configuration, it still forms ring broadcast storm or ring block, please open debug function of primary node MRPP, and used show MRPP statistics command to observe states of primary node and transfer node and statistics information is normal or not, and then sends results to our Technology Service Center.

Chapter 34 ULPP Configuration

34.1 Introduction to ULPP

Each ULPP group has two uplink ports, they are master port and slave port. The port may be a physical port or a port channel. The member ports of ULPP group have three states: Forwarding, Standby, Down. Normally, only one port at the forwarding state, the other port is blocked at the Standby state. When the master port has the link problem, the master port becomes down state, and the slave port is siwthed to forwarding state.



The using scene of ULPP

The above figure uses the double-uplink network, this is the typical application scene of ULPP. SwitchA goes up to SwitchD through SwitchB and SwitchC, port A1 and port A2 are the uplink ports. SwitchA configures ULPP, thereinto port A1 is set as the master port, port A2 is set as the slave port. When port A1 at forwarding state has the problem, switch the uplink at once, port A2 turns into forwarding state. After this, when recovering the master port, if the preemption mode is not configured, port A2 keeps the Forwarding state, port A1 turns into the Standby state.

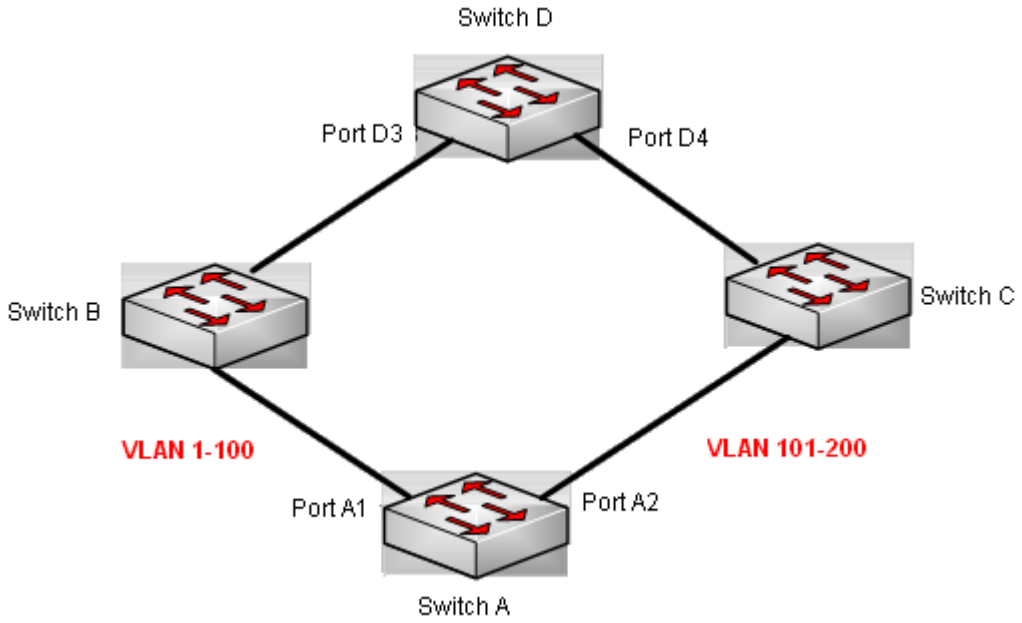
After the preemption mode is enabled, so as to the master port preempts the slave port when it recovered from the problem. For avoiding the frequent uplink switch caused by the abnormity problem, the preemption delay mechanism is imported, and it needs to wait for some times before the master port preempt the slave port. For keeping the continuance of the flows, the master port does not process to preempt by default, but turns into the Standby state.

When configuring ULPP, it needs to specify the VLAN which is protected by this ULPP group

through the method of MSTP instances, and ULPP does not provide the protection to other VLANs.

When the uplink switch is happennig, the primary forwarding entries of the device will not be applied to new topology in the network. In the figure, SwitchA configures ULPP, the portA1 as the master port at forwarding state, here the MAC address of PC is learned by Switch D from portD3. After this, portA1 has the problem, the traffic is switched to portA2 to be forwarded. If there is the data sent to PC by SwitchD, still the data will be forwarded from portD3, and will be losed. Therefore, when switching the uplink, the device of configuring ULPP needs to send the flush packets through the port which is switched to Forwarding state, and update MAC address tables and ARP tables of other devices in the network. ULPP respectively uses two kinds of flush packets to update the entries: the updated packets of MAC address and the deleted packets of ARP.

For making use of the bandwidth resource enough, ULPP can implement VLAN load balance through the configuration. As the picture illustrated, SwitchA configures two ULPP groups: portA1 is the master port and portA2 is the slave port in group1, portA2 is the master port and portA1 is the slave port in group2, the VLANs are protected by group1 and group2, they are 1-100 and 101-200. Here both portA1 and portA2 at the forwarding state, the master port and the slave port mutually backup, and respectively forward the packets of the different VLAN ranges. When portA1 has the problem, the traffic of VLAN 1-200 are forwarded by portA2. After this, when portA1 is recovering the normal state, portA2 forwards the data of VLAN 101-200 sequentially, but the data of VLAN 1-100 is switched to portA1 to forward.



VLAN load balance

34.2 ULPP Configuration Task List

Create ULPP group globally

Configure ULPP group

Show and debug the relating information of ULPP

1. Create ULPP group globally

Command	Expalnation
Global mode	
ulpp group <*integer*> **no ulpp group <*integer*>**	Configure and delete ULPP group globally.

2. Configure ULPP group

Command	Explanation
ULPP group configuration mode	
preemption mode **no preemption mode**	Configure the preemption mode of ULPP group. The no operation deletes the preemption mode.
preemption delay <*integer*> **no preemption delay**	Configure the preemption delay, the no operation restores the default value 30s.
control vlan <*integer*> **no control vlan**	Configure the sending control VLAN, no operation restores the default value 1.
protect vlan-reference-instance <*instance-list*> **no protect vlan-reference-instance <*instance-list*>**	Configure the protection VLANs, the no operation deletes the protection VLANs.
flush enable mac **flush disable mac**	Enable or disable sending the flush packets which update MAC address.
flush enable arp **flush disable arp**	Enable or disable sending the flush packets which delete ARP.
description <*string*> **no description**	Configure or delete ULPP group description.
Port mode	
ulpp control vlan <*vlan-list*> **no ulpp control vlan <*vlan-list*>**	Configure the receiving control VLANs, no operation restores the default value 1.
ulpp flush enable mac **ulpp flush disable mac**	Enable or disable receiving the flush packets which update the MAC address.
ulpp flush enable arp **ulpp flush disable arp**	Enable or disable receiving the flush packets which delete ARP.
ulpp group <*integer*> master **no ulpp group <*integer*> master**	Configure or delete the master port of ULPP group.

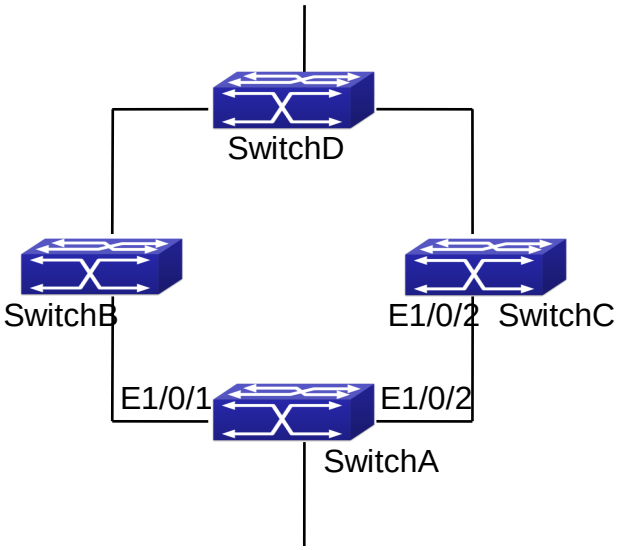
ulpp group <*integer*> slave **no ulpp group <*integer*> slave**	Configure or delete the slave port of ULPP group.

3. Show and debug the relating information of ULPP

Command	Explanation
Admin mode	
show ulpp group [group-id]	Show the configuration information of the configured ULPP group.
show ulpp flush counter interface {ethernet <IFNAME> | <IFNAME>}	Show the statistic information of the flush packets.
show ulpp flush-receive-port	Show flush type and control VLAN received by the port.
clear ulpp flush counter interface <*name*>	Clear the statistic information of the flush packets.
debug ulpp flush {send | receive} interface <*name*> **no debug ulpp flush {send | receive} interface <*name*>**	Show the information of the receiving and sending flush packets, the no operation disables the shown information.
debug ulpp flush content interface <*name*> **no debug ulpp flush content interface <*name*>**	Show the contents of the received flush packets, the no operation disables the showing.
debug ulpp error **no debug ulpp error**	Show the error information of ULPP, the no operation disables the showing.
debug ulpp event **no debug ulpp event**	Show the event information of ULPP, the no operation disables the showing.

34.3 ULPP Typical Examples

34.3.1 ULPP Typical Example1



ULPP typical example1

The above topology is the typical application environment of ULPP protocol.

SwitchA has two uplinks, they are SwitchB and SwitchC. When any protocols are not enabled, this topology forms a ring. For avoiding the loopback, SwitchA can configure ULPP protocol, the master port and the slave port of ULPP group. When both master port and slave port are up, the slave port will be set as standby state and will not forward the data packets. When the master port is down, the slave port will be set as forwarding state and switch to the uplink. SwitchB and SwitchC can enable the command that receives the flush packets, it is used to associate with ULPP protocol running of SwitchA to switch the uplink immediately and reduce the switch delay.

When configuring ULPP protocol of SwitchA, first, create a ULPP group and configure the protection VLAN of this group as vlan10, then configure interface Ethernet 1/0/1 as the master port, interface Ethernet 1/0/2 as the slave port, the control VLAN as 10. SwitchB and SwitchC configure the flush packets that receive ULPP.

SwitchA configuration task list:

Switch(Config)#vlan 10

Switch(Config-vlan10)#switchport interface ethernet 1/0/1; 1/0/2

Switch(Config-vlan10)#exit

Switch(Config)#spanning-tree mst configuration

Switch(Config-Mstp-Region)#instance 1 vlan 10

Switch(Config-Mstp-Region)#exit

Switch(Config)#ulpp group 1

Switch(ulpp-group-1)#protect vlan-reference-instance 1

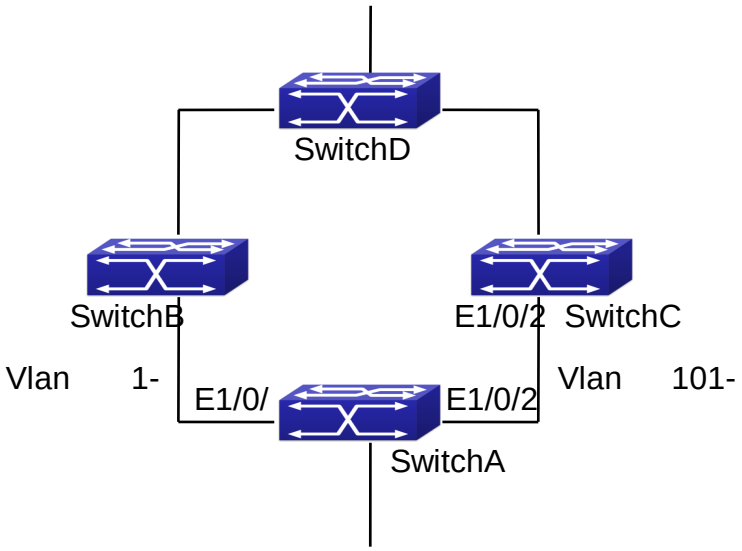
Switch(ulpp-group-1)#control vlan 10

Switch(ulpp-group-1)#exit
Switch(Config)#interface ethernet 1/0/1
Switch(config-If-Ethernet1/0/1)# ulpp group 1 master
Switch(config-If-Ethernet1/0/1)#exit
Switch(Config)#interface Ethernet 1/0/2
Switch(config-If-Ethernet1/0/2)# ulpp group 1 slave
Switch(config-If-Ethernet1/0/2)#exit

SwitchB configuration task list:
Switch(Config)#vlan 10
Switch(Config-vlan10)#switchport interface ethernet 1/0/1
Switch(Config-vlan10)#exit
Switch(Config)#interface ethernet 1/0/1
Switch(config-If-Ethernet1/0/1)# ulpp flush enable mac
Switch(config-If-Ethernet1/0/1)# ulpp flush enable arp
Switch(config-If-Ethernet1/0/1)# ulpp control vlan 10

SwitchC configuration task list:
Switch(Config)#vlan 10
Switch(Config-vlan10)#switchport interface ethernet 1/0/2
Switch(Config-vlan10)#exit
Switch(Config)#interface ethernet 1/0/2
Switch(config-If-Ethernet1/0/2)# ulpp flush enable mac
Switch(config-If-Ethernet1/0/2)# ulpp flush enable arp
Switch(config-If-Ethernet1/0/2)# ulpp control vlan 10

34.3.2 ULPP Typical Example2



ULPP typical example2

ULPP can implement the VLAN-based load balance. As the picture illustrated, SwitchA configures two ULPP groups: port E1/0/1 is the master port and port 1/0/2 is the slave port in group1, port 1/0/2 is the master port and port 1/0/1 is the slave port in group2. The VLANs protected by group1 are 1-100 and by group2 are 101-200. Here both port E1/0/1 and port E1/0/2 at the forwarding state, the master port and the slave port mutually backup, respectively forward the packets of different VLAN ranges. When port E1/0/1 has the problem, the traffic of VLAN 1-200 are forwarded by port E1/0/2. When port E1/0/1 is recovering the normal state, still port E1/0/2 forwards the data of VLAN 101-200, the data of VLAN 1-100 are switched to port E1/0/1 to forward.

SwitchA configuration task list:

```
Switch(Config)#spanning-tree mst configuration
Switch(Config-Mstp-Region)#instance 1 vlan 1-100
Switch(Config-Mstp-Region)#instance 2 vlan 101-200
Switch(Config-Mstp-Region)#exit
Switch(Config)#ulpp group 1
Switch(ulpp-group-1)#protect vlan-reference-instance 1
Switch(ulpp-group-1)#preemption mode
Switch(ulpp-group-1)#exit
Switch(Config)#ulpp group 2
Switch(ulpp-group-2)#protect vlan-reference-instance 2
Switch(ulpp-group-1)#preemption mode
Switch(ulpp-group-2)#exit
Switch(Config)#interface ethernet 1/0/1
Switch(config-If-Ethernet1/0/1)#switchport mode trunk
```

Switch(config-If-Ethernet1/0/1)#ulpp group 1 master
Switch(config-If-Ethernet1/0/1)#ulpp group 2 slave
Switch(config-If-Ethernet1/0/1)#exit
Switch(Config)#interface Ethernet 1/0/2
Switch(config-If-Ethernet1/0/2)#switchport mode trunk
Switch(config-If-Ethernet1/0/2)# ulpp group 1 slave
Switch(config-If-Ethernet1/0/2)# ulpp group 2 master
Switch(config-If-Ethernet1/0/2)#exit

SwitchB configuration task list:
Switch(Config)#interface ethernet 1/0/1
Switch(config-If-Ethernet1/0/1)#switchport mode trunk
Switch(config-If-Ethernet1/0/1)# ulpp flush enable mac
Switch(config-If-Ethernet1/0/1)# ulpp flush enable arp

SwitchC configuration task list:
Switch(Config)#interface ethernet 1/0/2
Switch(config-If-Ethernet1/0/2)# switchport mode trunk
Switch(config-If-Ethernet1/0/2)# ulpp flush enable mac
Switch(config-If-Ethernet1/0/2)# ulpp flush enable arp

34.4 ULPP Troubleshooting

At present, configuration of more than 2 multi-uplinks is allowed, but it may cause loopback, so is not recommended.
With the normal configuration, if the broadcast storm happen or the communication along the ring is broken, please enable the debug of ULPP, copy the debug information of 3 minutes and the configuration information, send them to our technical service center.

Chapter 35 ULSM Configuration

35.1 Introduction to ULSM

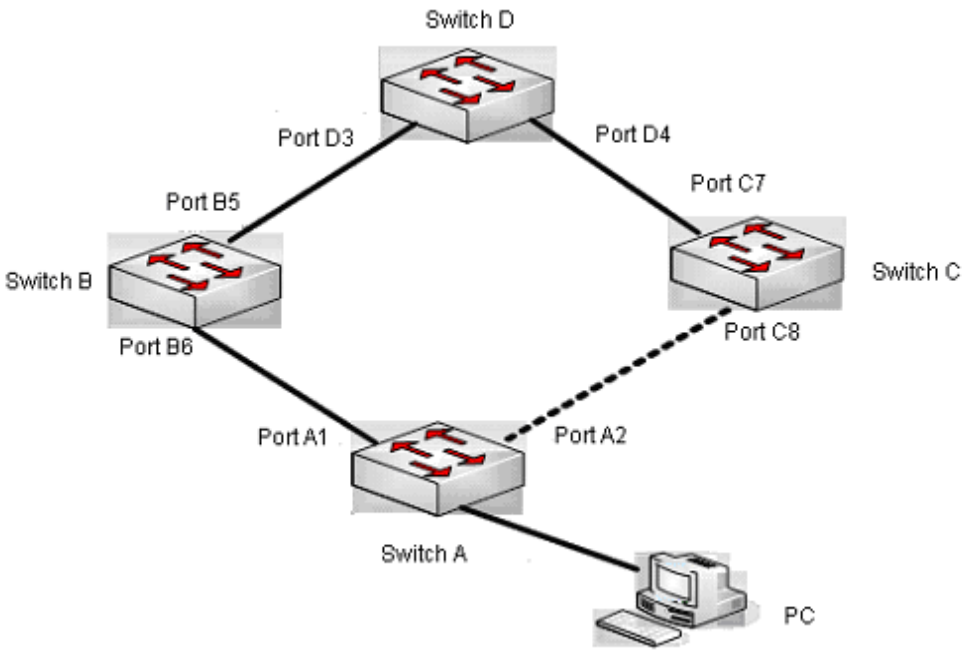
ULSM (Uplink State Monitor) is used to process the port state synchronization. Each ULSM group is made up of the uplink port and the downlink port, both the uplink port and the downlink port may be multiple. The port may be a physical port or a port channel, but it can not be a member port of a port channel, and each port only belongs to one ULSM group.

The uplink port is the monitored port of ULSM group. When all uplink ports are down or there is no uplink port in ULSM group, ULSM group state is down. ULSM group state is up as long as one uplink port is up.

The downlink port is the controlled port, its state changes along with Up/Down of ULSM group and is always the same with ULSM group state.

ULSM associates with ULPP to enable the downstream device to apperceive the link problem of the upstream device and process correctly. As the picture illustrated, SwitchA configures ULPP, here the traffic is forwarded by port A1. If the link between SwitchB and Switch D has the problem, SwitchA can not apperceive the problem of the upstream link and sequentially forward the traffic from port A1, cause traffic losing.

Configuring ULSM on SwitchB can solve the above problems. The steps are: set port B5 as the uplink port of ULSM group, port B6 as the downlink port. When the link between SwitchB and SwitchD has the problem, both the downlink port B6 and the state of ULSM group are down. It causes Switch A on which ULPP is configured to process uplink switchover and avoid the data dropped.



35.2 ULSM Configuration Task List

Create ULSM group globally
Configure ULSM group
Show and debug the relating information of ULSM

1. Create ULSM group globally

Command	explanation
Global mode	
ulsm group <group-id> **no ulsm group <*group-id*>**	Configure and delete ULSM group globally.

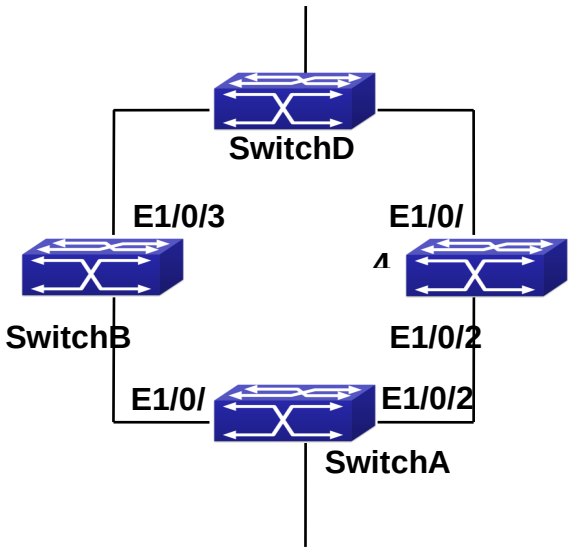
2. Configure ULSM group

Command	explanation
Port mode	
ulsm group <*group-id*> {uplink | downlink} **no ulsm group <*group-id*> {uplink | downlink}**	Configure the uplink/downlink port of ULSM group, the no command deletes the uplink/downlink port.

3. Show and debug the relating information of ULSM

Command	Explanation
Admin mode	
show ulsm group [group-id]	Show the configuration information of ULSM group.
debug ulsm event **no debug ulsm event**	Show the event information of ULSM, the no operation disables the shown information.

35.3 ULSM Typical Example



ULSM typical example

The above topology is the typical application environment which is used by ULSM and ULPP protocol.

ULSM is used to process the port state synchronization, its independent running is useless, so it usually associates with ULPP protocol to use. In the topology, SwitchA enables ULPP protocol, it is used to switch the uplink. SwitchB and SwitchC enable ULSM protocol to monitor whether the uplink is down. If it is down, then ULSM will execute the down operation for the downlink port to shutdown it, so ULPP protocol of Swtich A executes the relative operation of the uplink switchover.

SwitchA configuration task list:

Switch(Config)#spanning-tree mst configuration
Switch(Config-Mstp-Region)#instance 1 vlan 1
Switch(Config-Mstp-Region)#exit
Switch(Config)#ulpp group 1
Switch(ulpp-group-1)#protect vlan-reference-instance 1
Switch(ulpp-group-1)#exit
Switch(Config)#interface ethernet 1/0/1
Switch(config-If-Ethernet1/0/1)# ulpp group 1 master
Switch(config-If-Ethernet1/0/1)#exit
Switch(Config)#interface Ethernet 1/0/2
Switch(config-If-Ethernet1/0/2)# ulpp group 1 slave
Switch(config-If-Ethernet1/0/2)#exit

SwitchB configuration task list:

Switch(Config)#ulsm group 1

Switch(Config)#interface ethernet 1/0/1
Switch(config-If-Ethernet1/0/1)#ulsm group 1 downlink
Switch(config-If-Ethernet1/0/1)#exit
Switch(Config)#interface ethernet 1/0/3
Switch(config-If-Ethernet1/0/3)#ulsm group 1 uplink
Switch(config-If-Ethernet1/0/3)#exit

SwitchC configuration task list:
Switch(Config)#ulsm group 1
Switch(Config)#interface ethernet 1/0/2
Switch(config-If-Ethernet1/0/2)#ulsm group 1 downlink
Switch(config-If-Ethernet1/0/2)#exit
Switch(Config)#interface ethernet 1/0/4
Switch(config-If-Ethernet1/0/4)#ulsm group 1 uplink
Switch(config-If-Ethernet1/0/4)#exit

35.4 ULSM Troubleshooting

With the normal configuration, if the downlink port does not responds the down event of the uplink port, please enable the debug function of ULSM, copy the debug information of 3 minutes and the configuration information, and send them to our technical service center.

Chapter 36 Mirror Configuration

36.1 Introduction to Mirror

Mirror functions include port mirror function, flow mirror function.

Port mirror refers to the duplication of data frames sent/received on a port to another port. The duplicated port is referred to as mirror source port and the duplicating port is referred to as mirror destination port. A protocol analyzer (such as Sniffer) or a RMON monitor will be connected at mirror destination port to monitor and manage the network, and diagnose the problems in the network.

Flow mirror function means that the switch exactly copies the data frames received or by the specified rule of a port to another port. The flow mirror will take effect only the specified rule is permit.

A chassis switch supports at most 4 mirror destination ports, each boardcard allows a source or destination port of a mirror session. At present, each box switch can set many mirror sessions. There is no limitation on mirror source ports, one port or several ports is allowed. When there are more than one source ports, they can be in the same VLAN or in different VLAN. The source port and destination port can be in different VLAN.

36.2 Mirror Configuration Task List

1. Specify mirror destination port
2. Specify mirror source port
3. Specify flow mirror source

1. Specify mirror destination port

Command	Explanation
Global mode	
monitor session <*session*> destination interface <*interface-number*> **no monitor session <*session*> destination interface <*interface-number*>**	Specifies mirror destination port; the no command deletes mirror destination source port.

2. Specify mirror source port

Command	Explanation
Global mode	
monitor session <*session*> source	Specifies mirror source port; the no command

{interface <*interface-list*>} {rx| tx| both} **no monitor session <*session*> source {interface <*interface-list*>}**	deletes mirror source port.

3. Specify flow mirror source

Command	Explanation
Global mode	
monitor session <*session*> source {interface <*interface-list*>} access-group <*num*> {rx|tx|both} **no monitor session <*session*> source {interface <*interface-list*>} access-group <*num*>**	Specifies flow mirror source port and apply rule; the no command deletes flow mirror source port.

36.3 Mirror Examples

Example:

The requirement of the configurations is shown as below: to monitor at interface 1 the data frames sent out by interface 9 and received from interface 7, and the data frames received by interface 15 and matched by rule 120(The source IP address is 1.2.3.4 and the destination IP address is 5.6.7.8).

Configuration guidelines:

1. Configure interface 1 to be a mirror destination interface.
2. Configure the interface 7 ingress and interface 9 egress to be mirrored source.
3. Configure access list 120.
4. Configure access 120 to binding interface 15 ingress.

Configuration procedure is as follows:

Switch(config)#monitor session 4 destination interface ethernet 1/0/1
Switch(config)#monitor session 4 source interface ethernet 1/0/7 rx
Switch(config)#monitor session 4 source interface ethernet 1/0/9 tx
Switch(config)#access-list 120 permit tcp 1.2.3.4 0.0.0.255 5.6.7.8 0.0.0.255
Switch(config)#monitor session 4 source interface ethernet 1/0/15 access-list 120 rx

36.4 Device Mirror Troubleshooting

If problems occur on configuring port mirroring, please check the following first for causes:

Whether the mirror destination port is a member of a TRUNK group or not, if yes, modify the TRUNK group.

If the throughput of mirror destination port is smaller than the total throughput of mirror source port(s), the destination port will not be able to duplicate all source port traffic; please decrease the number of source ports, duplicate traffic for one direction only or choose a port with greater throughput as the destination port. Mirror destination port can not be pulled into Isolate vlan, or will affect mirror between VLAN.

Chapter 37 RSPAN Configuration

37.1 Introduction to RSPAN

Port mirroring refers to the duplication of data frames sent/received on a port to another port. The duplicated port is referred to as mirror source port and the duplicating port is referred to as mirror destination port. It is more convenience for network administrator to monitor and manage the network and diagnostic after the mirroring function achieved. But it only used for such instance that the mirror source port and the mirror destination ports are located in the same switch.

RSPAN (remote switched port analyzer) refers to remote port mirroring. It eliminates the limitation that the source port and the destination port must be located on the same switch. This feature makes it possible for the source port and the destination port to be located on different devices in the network, and facilitates the network administrator to manage remote switches. It can't forward traffic flows on remote mirror VLAN.

There are three types of switches with the RSPAN enabled:

1. Source switch: The switch to which the monitored port belongs. The source switch copies the mirrored traffic flows to the Remote VLAN, and then through Layer 2 forwarding, the mirrored flows are sent to an intermediate switch or destination switch.
2. Intermediate switch: Switches between the source switch and destination switch on the network. Intermediate switch forwards mirrored flows to the next intermediate switch or the destination switch. Circumstances can occur where no intermediate switch is present, if a direct connection exists between the source and destination switches.
3. Destination switch: The switch to which the destination port for remote mirroring belongs. It forwards mirrored flows it received from the Remote VLAN to the monitoring device through the destination port.

When configuring the RSPAN mirroring of the source switch, reflector port mode or destination mirror port mode can be selected. The destination switch will redirect all the data frames in the RSPAN VLAN to the RSPAN destination port. For RSPAN mirroring, normal mode and advanced mode can be chosen, normal is introduced by default and fit the normal user. The advanced mode fit the advanced user.

1. Advanced mode: To redirect data frames in RSPAN VLAN to the RSPAN destination port, the intermediary and destination devices should support the redirection of flow.
2. Normal mode: To configure the RSPAN destination port in the RSPAN VLAN. Thus, datagrams in the RSPAN VLAN will be broadcasted to the destination port. In this mode, the destination port should be in RSPAN VLAN, and the source port should not be configured for broadcasting storm control. TRUNK ports should be configured carefully in order not to forward RSPAN datagrams to external networks. The normal mode has the benefit of easy

configuration, and reduced system resources.

To be noticed: Normal mode is introduced by default. When using the normal mode, datagrams with reserved MAC addresses cannot be broadcasted.

The number of the source mirror ports is not limited, and can be one or more. Multiple source ports are not restricted to be in the same VLAN. The destination port and the source ports can be in different VLAN.

For configuration of RSPAN, a dedicated RSPAN VLAN should be configured first for carrying the RSPAN datagrams. The default VLAN, dynamic VLAN, private VLAN, multicast VLAN, and the layer 3 interface enabled VLAN cannot be configured as the RSPAN VLAN. The reflector port must belong to the RSPAN VLAN. The destination port should be connected to the Monitor and the configured as access port or the TRUNK port. The RSPAN reflector port will be working dedicatedly for mirroring, when a port is configured as a reflector port, it will discards all the existing connections to the remote peer, disable configurations related to loopback interfaces, and stop forwarding datagram. Connectivity between the source and destination switch for Remote VLAN, should be made sure by configuration.

To be noticed:

1. Layer 3 interfaces related to RSPAN VLAN should not be configured on the source, intermediate, and the destination switches, or the mirrored datagrams may be discarded.
2. For the source and intermediate switches in the RSPAN connections, the native VLAN of TRUNK port cannot be configured as the RSPAN VLAN, Otherwise the RSPAN tag will be disposed before reaching the destination switches.
3. The source port, in access or trunk mode, should not be added to RSPAN VLAN if advanced RSPAN mode is chosen. When the reflector port is used for a inter-card mirroring of CPU TX data, it must be configured as TRUNK port and allows the RSPAN VLAN data passing, the Native VLAN should not be configured as RSPAN VLAN.
4. When configuring the remote mirroring function, the network bandwidth should be considered in order to carry the network flow and the mirrored flow.

Keywards:

RSPAN: Remote Switched Port Analyzer

RSPAN VLAN: Dedicated VLAN for RSPAN

RSPAN Tag: The VLAN tag which is attached to MTP of the RSPAN datagrams.

Reflector Port: The local mirroring port between the RSPAN source and destination ports, which is not directly connected to the intermediate switches.

37.2 RSPAN Configuration Task List

Configure RSPAN VLAN

Configure mirror source port

Configure mirror destination port

Configure reflector port

Configure remote VLAN of mirror group

1. Configure RSPAN VLAN

Command	Explanation
VLAN Configuration Mode	
remote-span **no remote-span**	To configure the specified VLAN as RSPAN VLAN. The no command will remove the configuration of RSPAN VLAN.

2. Configure mirror source port

Command	Explanation
Global Mode	
monitor session <*session*> source {interface <*interface-list*>} {rx| tx| both} **no monitor session <*session*> source {interface <*interface-list*>}**	To configure mirror source port; The no command deletes the mirror source port.

3. Configure mirror destination port

Command	Explanation
Global Mode	
monitor session <*session*> destination interface <*interface-number*> **no monitor session <*session*> destination interface <*interface-number*>**	To configure mirror destination interface; The no command deletes the mirror destination port.

4. Configure reflector port

Command	Explanation
Global Mode	
monitor session <session> reflector-port <interface-number> **no monitor session <session> reflector-port**	To configure the interface to reflector port; The no command deletes the reflector port.

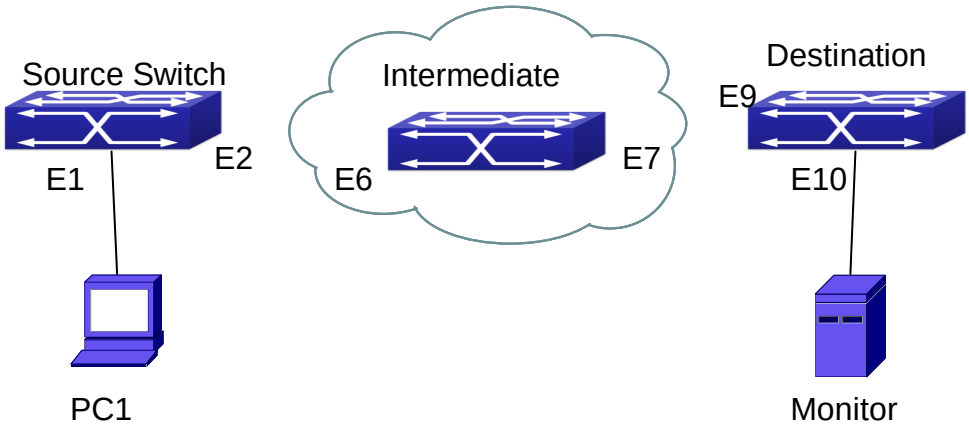
5. Configure remote VLAN of mirror group

Command	Explanation
Global Mode	
monitor session <*session*> **remote vlan <*vid*>** **no monitor session <session> remote vlan**	To configure remote VLAN of mirror group, the no command deletes the remote VLAN of mirror group.

37.3 Typical Examples of RSPAN

Before RSPAN is invented, network administrators had to connect their PCs directly to the switches, in order to check the statistics of the network.

However, with the help of RSPAN, the network administrators can configure and supervise the switches remotely, which brings more efficiency. The figure below shows a sample application of RSPAN.



RSPAN Application Sample

Two configuration solutions can be chosen for RSPAN: the first is without reflector port, and the other is with reflector port. For the first one, only one fixed port can be connected to the intermediate switch. However, no reflector port has to be configured. This maximizes the usage of witch ports. For the latter one, the port connected to the intermediate switch is not fixed. Datagrams can be broadcasted in the RSPAN VLAN through the loopback, which is much more flexible.

The normal mode configuration is show as below:

Solution 1:

Source switch:

Interface ethernet 1/0/1 is the source port for mirroring.

Interface ethernet 1/0/2 is the destination port which is connected to the intermediate switch.

RSPAN VLAN is 5.

Switch(config)#vlan 5
Switch(Config-Vlan5)#remote-span
Switch(Config-Vlan5)#exit
Switch(config)#interface ethernet 1/0/2
Switch(Config-If-Ethernet1/0/2)#switchport mode trunk
Switch(Config-If-Ethernet1/0/2)#exit
Switch(config)#monitor session 1 source interface ethernet1/0/1 rx
Switch(config)#monitor session 1 destination interface ethernet1/0/2
Switch(config)#monitor session 1 remote vlan 5

Intermediate switch:
Interface ethernet1/0/6 is the source port which is connected to the source switch.
Interface ethernet1/0/7 is the destination port which is connected to the intermediate switch. The native VLAN of this port cannot be configured as RSPAN VLAN, or the mirrored data may not be carried by the destination switch.
RSPAN VLAN is 5.

Switch(config)#vlan 5
Switch(Config-Vlan5)#remote-span
Switch(Config-Vlan5)#exit
Switch(config)#interface ethernet 1/0/6-7
Switch(Config-If-Port-Range)#switchport mode trunk
Switch(Config-If-Port-Range)#exit

Destination switch:
Interface ethernet1/0/9 is the source port, which is connected to the source switch.
Interface ethernet1/0/10 is the destination port which is connected to the monitor. This port is required to be configured as an access port, and belong to the RSPAN VLAN.
RSPAN VLAN is 5.

Switch(config)#vlan 5
Switch(Config-Vlan5)#remote-span
Switch(Config-Vlan5)#exit
Switch(config)#interface ethernet 1/0/9
Switch(Config-If-Ethernet1/0/9)#switchport mode trunk
Switch(Config-If-Ethernet1/0/9)#exit
Switch(config)#interface ethernet 1/0/10
Switch(Config-If-Ethernet1/0/10)#switchport access vlan 5

Switch(Config-If-Ethernet1/0/10)#exit

Solution 2:
Source switch:
Interface ethernet 1/0/1 is the source port.
Interface ethernet 1/0/2 is the TRUNK port, which is connected to the intermediate switch. The native VLAN should not be a RSPAN VLAN.
Interface Ethernet 1/0/3 is a reflector port. The reflector port belongs the RSPAN VLAN, it is access port or TRUNK port of the RSPAN VLAN.
RSPAN VLAN is 5.

Switch(config)#vlan 5
Switch(Config-Vlan5)#remote-span
Switch(Config-Vlan5)#exit
Switch(config)#interface ethernet 1/0/2
Switch(Config-If-Ethernet1/0/2)#switchport mode trunk
Switch(Config-If-Ethernet1/0/2)#exit
Switch(config)#interface ethernet 1/0/3
Switch(Config-If-Ethernet1/0/3)#switchport mode trunk
Switch(Config-If-Ethernet1/0/3)#exit
Switch(config)#monitor session 1 source interface ethernet1/0/1 rx
Switch(config)#monitor session 1 reflector-port ethernet1/0/3
Switch(config)#monitor session 1 remote vlan 5

Intermediate switch:
Interface ethernet1/0/6 is the source port which is connected to the source switch.
Interface ethernet1/0/7 is the destination port which is connected to the destination switch. The native VLAN of the port should not be configured as RSPAN VLAN, or the mirrored data may not be carried by the destination switch.
RSPAN VLAN is 5.

Switch(config)#vlan 5
Switch(Config-Vlan5)#remote-span
Switch(Config-Vlan5)#exit
Switch(config)#interface ethernet 1/0/6-7
Switch(Config-If-Port-Range)#switchport mode trunk
Switch(Config-If-Port-Range)#exit

Destination switch:

Interface ethernet1/0/9 is the source port which is connected to the source switch.
Interface ethernet1/0/10 is the destination port which is connected to the monitor. This port is required to be configured as an access port, and belong to the RSPAN VLAN.
RSPAN VLAN is 5.

```
Switch(config)#vlan 5
Switch(Config-Vlan5)#remote-span
Switch(Config-Vlan5)#exit
Switch(config)#interface ethernet 1/0/9
Switch(Config-If-Ethernet1/0/9)#switchport mode trunk
Switch(Config-If-Ethernet1/0/9)#exit
Switch(config)#interface ethernet 1/0/10
Switch(Config-If-Ethernet1/0/10)#switchport access vlan 5
Switch(Config-If-Ethernet1/0/10)#exit
```

37.4 RSPAN Troubleshooting

Due to the following reasons, RSPAN may not function:
Whether the destination mirror port is a member of the Port-channel group. If so, please change the Port-channel group configuration;
The throughput the destination port is less than the total throughput of the source mirror ports. If so, the destination cannot catch all the datagrams from every source ports. To solve the problem, please reduce the number of the source ports, or mirror only single direction data flow, or choose some other port with higher capacity as the destination port.
Between the source switch and the intermediate switch, whether the native VLAN of the TRUNK ports is configured as RSPAN VLAN. If so, please change the native VLAN for the TRUNK ports.

Chapter 38 sFlow Configuration

38.1 Introduction to sFlow

The sFlow (RFC 3176) is a protocol based on standard network export and used on monitoring the network traffic information developed by the InMon Company. The monitored switch or router sends date to the client analyzer through its main operations such as sampling and statistic, then the analyzer will analyze according to the user requirements so to monitor the network.

A sFlow monitor system includes: sFlow proxy, central data collector and sFlow analyzer. The sFlow proxy collects data from the switch using sampling technology. The sFlow collector is for formatting the sample data statistic which is to be forwarded to the sFlow analyzer which will analyze the sample data and perform corresponding measure according to the result. Our switch here acts as the proxy and central data collector in the sFlow system.

We have achieved data sampling and statistic targeting physical port.

Our data sample includes the IPv4 and IPv6 packets. Extensions of other types are not supported so far. As for non IPv4 and IPv6 packet, the unify HEADER mode will be adopted following the requirements in RFC3176, copying the head information of the packet based on analyzing the type of its protocol.

The latest sFlow protocol presented by InMon Company is the version 5. Since it is the version 4 which is realized in the RFC3176, version conflict might exist in some case such as the structure and the packet format. This is because the version 5 has not become the official protocol, so, in order to be compatible with current applications, we will continue to follow the RFC3176.

38.2 sFlow Configuration Task List

1. Configure sFlow Collector address

Command	Explanation
Global mode and Port Mode	
sflow destination <collector-address> [<collector-port>] **no sflow destination**	Configure the IP address and port number of the host in which the sFlow analysis software is installed. As for the ports, if IP address is configured on the port, the port configuration will be applied, or else will be applied the global configuration. The "**no sflow destination**" command restores to the default port value and deletes the IP address.

2. Configure the sFlow proxy address

Command	Explanation
Global Mode	
sflow agent-address <collector-address> **no sflow agent-address**	Configure the source IP address applied by the sFlow proxy; the “no” form of the command deletes this address.

3. Configure the sFlow proxy priority

Command	Explanation
Global Mode	
sflow priority <priority-vlaue> **no sflow priority**	Configure the priority when sFlow receives packet from the hardware; the “**no sflow priority**” command restores to the default

4. Configure the packet head length copied by sFlow

Command	Explanation
Port Mode	
sflow header-len <*length-vlaue*> **no sflow header-len**	Configure the length of the packet data head copied in the sFlow data sampling; the “no” form of this command restores to the default value.

5. Configure the max data head length of the sFlow packet

Command	Explanation
Port Mode	
sflow data-len <length-vlaue> **no sflow data-len**	Configure the max length of the data packet in sFlow; the “no” form of this command restores to the default.

6. Configure the sampling rate value

Command	Explanation
Port Mode	
sflow rate {input <input-rate> | output <output-rate >} **no sflow rate [input | output]**	Configure the sampling rate when sFlow performing hardware sampling. The “no” command deletes the rate value.

7. Configure the sFlow statistic sampling interval

Command	Explanation
Port Mode	
sflow counter-interval <*interval-vlaue*> **no sflow counter-interval**	Configure the max interval when sFlow performing statistic sampling. The “no” form of this command deletes

8. Configure the analyzer used by sFlow

Command	Explanation
Global Mode	
sflow analyzer sflowtrend **no sflow analyzer sflowtrend**	Configure the analyzer used by sFlow, the no command deletes the analyzer.

38.3 sFlow Examples

sFlow configuration topology

As shown in the figure, sFlow sampling is enabled on the port 1/0/1 and 1/0/2 of the switch. Assume the sFlow analysis software is installed on the PC with the address of 192.168.1.200. The address of the layer 3 interface on the SwitchA connected with PC is 192.168.1.100. A loopback interface with the address of 10.1.144.2 is configured on the SwitchA. sFlow configuration is as follows:

Configuration procedure is as follows:

Switch#config

Switch (config)#sflow ageng-address 10.1.144.2

Switch (config)#sflow destination 192.168.1.200

Switch (config)#sflow priority 1

Switch (config)# interface ethernet1/0/1

Switch (Config-If-Ethernet1/0/1)#sflow rate input 10000

Switch (Config-If-Ethernet1/0/1)#sflow rate output 10000

Switch (Config-If-Ethernet1/0/1)#sflow counter-interval 20

Switch (Config-If-Ethernet1/0/1)#exit

Switch (config)# interface ethernet1/0/2
Switch (Config-If-Ethernet1/0/2)#sflow rate input 20000
Switch (Config-If-Ethernet1/0/2)#sflow rate output 20000
Switch (Config-If-Ethernet1/0/2)#sflow counter-interval 40

38.4 sFlow Troubleshooting

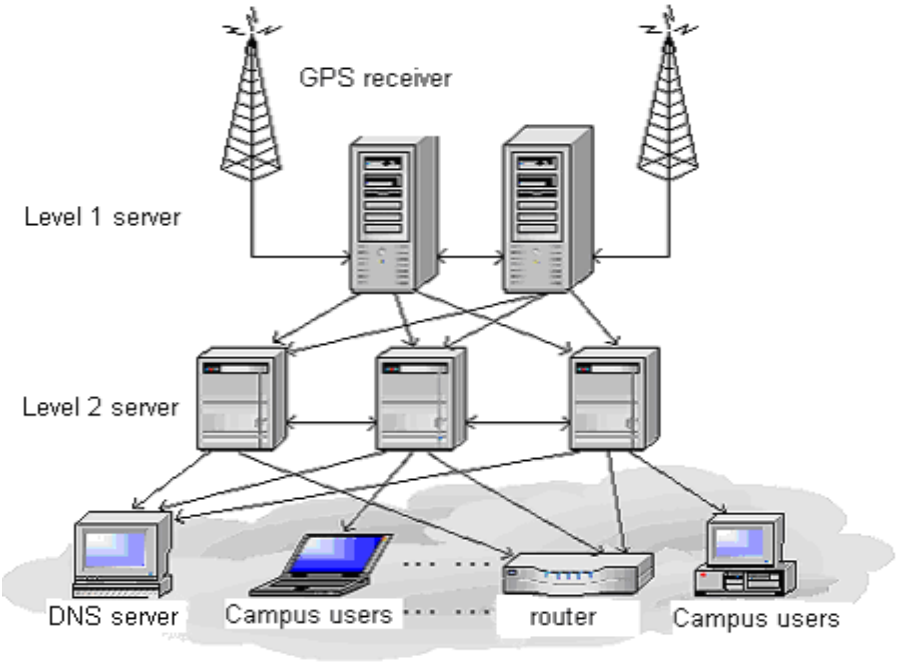
In configuring and using sFlow, the sFlow server may fail to run properly due to physical connection failure, wrong configuration, etc. The user should ensure the following:
Ensure the physical connection is correct
Guarantee the address of the sFlow analyzer configured under global or port mode is accessible.
If traffic sampling is required, the sampling rate of the interface must be configured
If statistic sampling is required, the statistic sampling interval of the interface must be configured
If the examination remains unsolved, please contact with the technical service center of our company.

Chapter 39 SNTP Configuration

39.1 Introduction to SNTP

The Network Time Protocol (NTP) is widely used for clock synchronization for global computers connected to the Internet. NTP can assess packet sending/receiving delay in the network, and estimate the computer's clock deviation independently, so as to achieve high accuracy in network computer clocking. In most positions, NTP can provide accuracy from 1 to 50ms according to the characteristics of the synchronization source and network route.

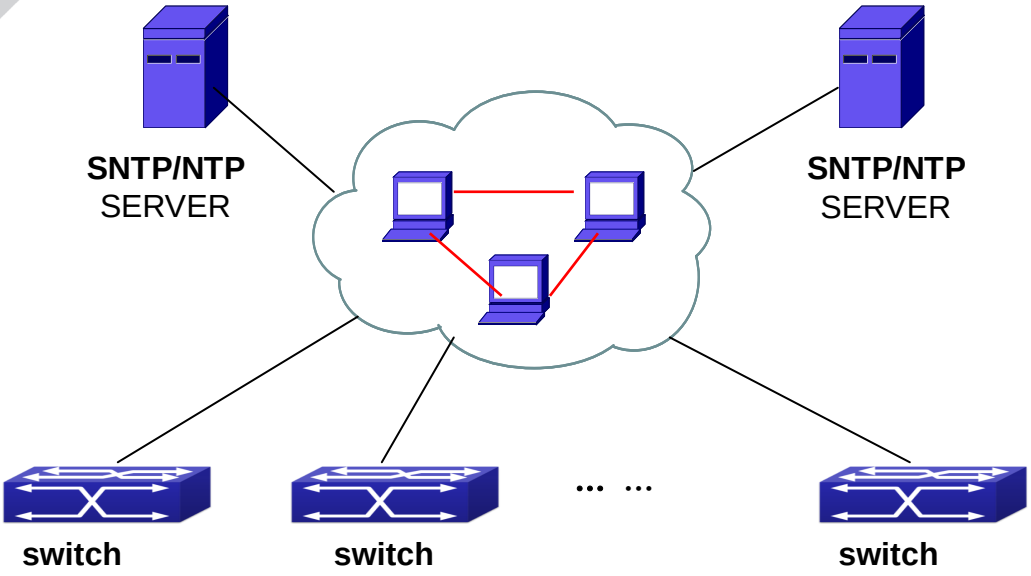
Simple Network Time Protocol (SNTP) is the simplified version of NTP, removing the complex algorithm of NTP. SNTP is used for hosts who do not require full NTP functions; it is a subset of NTP. It is common practice to synchronize the clocks of several hosts in local area network with other NTP hosts through the Internet, and use those hosts to provide time synchronization service for other clients in LAN. The figure below depicts a NTP/SNTP application network topology, where SNTP mainly works between second level servers and various terminals since such scenarios do not require very high time accuracy, and the accuracy of SNTP (1 to 50 ms) is usually sufficient for those services.



Working Scenario

Switch implements SNTPv4 and supports SNTP client unicast as described in RFC2030; SNTP client multicast and unicast are not supported, nor is the SNTP server function.

39.2 Typical Examples of SNTP Configuration



Typical SNTP Configuration

All switches in the autonomous zone are required to perform time synchronization, which is done through two redundant SNTP/NTP servers. For time to be synchronized, the network must be properly configured. There should be reachable route between any switch and the two SNTP/NTP servers.

Example: Assume the IP addresses of the SNTP/NTP servers are 10.1.1.1 and 20.1.1.1, respectively, and SNTP/NTP server function (such as NTP master) is enabled, then configurations for any switch should like the following:

Switch#config

Switch(config)#sntp server 10.1.1.1

Chapter 40 NTP Function Configuration

40.1 Introduction to NTP Function

The NTP (Network Time Protocol) synchronizes timekeeping spans WAN and LAN among distributed time servers and clients, it can get millisecond precision. The introduction of event, state, transmit function and action are defined in RFC-1305.

The purpose of using NTP is to keep consistent timekeeping among all clock-dependent devices within the network so that the devices can provide diverse applications based on the consistent time.

For a local system running NTP, its time can be synchronized by other reference sources and can be used as a reference source to synchronize other clocks, also can synchronize each other by transmit NTP packets.

40.2 NTP Function Configuration Task List

1. To enable NTP function
2. To configure NTP server function
3. To configure the max number of broadcast or multicast servers supported by the NTP client
4. To configure time zone
5. To configure NTP access control list
6. To configure NTP authentication
7. To specified some interface as NTP broadcast/multicast client interface
8. To configure some interface can't receive NTP packets
9. Display information
10. Debug

1. To enable NTP function

Command	Explication
Global Mode	
ntp enable **ntp disable**	To enable or disable NTP function.

2. To configure NTP server function

Command	Explication
Global Mode	
ntp server {<ip-address> | <ipv6-address>} [version <version_no>] [key	To enable the specified time server of time source.

<key-id>] **no ntp server {<ip-address> | <ipv6-address>}**	

3. To configure the max number of broadcast or multicast servers supported by the NTP client

Command	Explication
Global Mode	
ntp broadcast server count <number> **no ntp broadcast server count**	Set the max number of broadcast or multicast servers supported by the NTP client. The no operation will cancel the configuration and restore the default value.

4. To configure time zone

Command	Explication
Global Mode	
clock timezone WORD {add | subtract} <0-23> [<0-59>] **no clock timezone WORD**	This command configures timezone in global mode, the no command deletes the configured timezone.

5. To configure NTP access control list

Command	Explication
Global Mode	
ntp access-group server <*acl*> **no ntp access-group server < *acl*>**	To configure NTP server access control list.

6. To configure NTP authentication

Command	Explication
Global Mode	
ntp authenticate **no ntp authenticate**	To enable NTP authentication function.
ntp authentication-key <*key-id*> md5 <*value*> **no ntp authentication-key <*key-id*>**	To configure authentication key for NTP authentication.
ntp trusted-key <*key-id*> **no ntp trusted-key <*key-id*>**	To configure trusted key.

7. To specified some interface as NTP broadcast/multicast client interface

Command	Explication
Interface Configuration Mode	
ntp broadcast client **no ntp broadcast client**	To configure specified interface to receive NTP broadcast packets.
ntp multicast client **no ntp multicast client**	To configure specified interface to receive NTP multicast packets.
ntp ipv6 multicast client **no ntp ipv6 multicast client**	To configure specified interface to receive IPv6 NTP multicast packets.

8. To configure some interface can't receive NTP packets

Command	Explication
Interface Configuration Mode	
ntp disable **no ntp disable**	To disable the NTP function.

9. Display information

Command	Explication
Admin Mode	
show ntp status	To display the state of time synchronize.
show ntp session [*<ip-address>* | *<ipv6-address>*]	To display the information of NTP session.

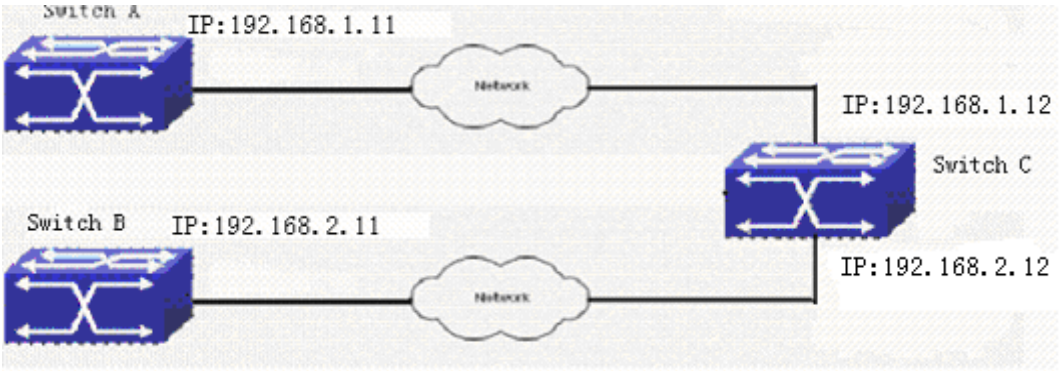
10. Debug

Command	Explication
Admin Mode	
debug ntp authentication **no debug ntp authentication**	To enable debug switch of NTP authentication.
debug ntp packets [send | receive] **no debug ntp packets [send | receive]**	To enable debug switch of NTP packet information.
debug ntp adjust **no debug ntp adjust**	To enable debug switch of time update information.

debug ntp sync **no debug ntp sync**	To enable debug switch of time synchronize information.
debug ntp events **no debug ntp events**	To enable debug switch of NTP event information.

40.3 Typical Examples of NTP Function

A client switch wanted to synchronize time with time server in network, there is two time server in network, the one is used as host, the other is used as standby, the connection and configuration as follows (Switch A and Switch B are the switch or route which support NTP server):



The configuration of Switch C is as follows: (Switch A and Switch B may have the different command because of different companies, we not explain there, our switches are not support NTP server at present)

Switch C:

Switch(config)#ntp enable

Switch(config)#interface vlan 1

Switch(Config-if-Vlan1)#ip address 192.168.1.12 255.255.255.0

Switch(config)#interface vlan 2

Switch(Config-if-Vlan1)#ip address 192.168.2.12 255.255.255.0

Switch(config)#ntp server 192.168.1.11

Switch(config)#ntp server 192.168.2.11

40.4 NTP Function Troubleshooting

In configuration procedures, if there is error occurred, the system can give out the debug information.

The NTP function disables by default, the show command can be used to display current

configuration. If the configuration is right please use debug every relative debugging command and display specific information in procedure, and the function is configured right or not, you can also use show command to display the NTP running information, any questions please send the recorded message to the technical service center.

Chapter 41 DNSv4/v6 Configuration

41.1 Introduction to DNS

DNS (Domain Name System) is a distributed database used by TCP/IP applications to translate domain names into corresponding IPv4/IPv6 addresses. With DNS, you can use easy-to-remember and signification domain names in some applications and let the DNS server translate them into correct IPv4/IPv6 addresses.

There are two types of DNS services, static and dynamic, which supplement each other in application. Each time the DNS server receives a name query it checks its static DNS database first before looking up the dynamic DNS database. Some frequently used addresses can be put in the static DNS database, the reduction the searching time in the dynamic DNS database would increase efficiency. The static domain name resolution means setting up mappings between domain names and IPv4/IPv6 addresses. IPv4/IPv6 addresses of the corresponding domain names can be found in the static DNS database when you use some applications. Dynamic domain name resolution is implemented by querying the DNS server. A user program sends a name query to the resolver in the DNS client when users want to use some applications with domain name, the DNS resolver looks up the local domain name cache for a match. If a match is found, it sends the corresponding IPv4/IPv6 address back to the switch. If no match is found, it sends a query to a higher DNS server. This process continues until a result, whether success or failure, is returned.

The Domain Name System (DNS) is a hierarchical naming system for computers, services, or any resource participating in the Internet. It associates various information with domain names assigned to such participants. Most importantly, it translates humanly meaningful domain names to the numerical (binary) identifiers associated with networking equipment for the purpose of locating and addressing these devices world-wide. An often used analogy to explain the Domain Name System is that it serves as the "phone book" for the Internet by translating human-friendly computer hostnames into IP addresses.

The Domain Name System makes it possible to assign domain names to groups of Internet users in a meaningful way, independent of each user's physical location. Because of this, World-Wide Web (WWW) hyperlinks and Internet contact information can remain consistent and constant even if the current Internet routing arrangements change or the participant uses a mobile device. Internet domain names are easier to remember than IP addresses such as 208.77.188.166(IPv4) or 2001:db8:1f70::999:de8:7648:6e8 (IPv6). People take advantage of this when they recite meaningful URLs and e-mail addresses without having to know how the machine will actually locate them.

The Domain Name System distributes the responsibility for assigning domain names and mapping them to Internet Protocol (IP) networks by designating authoritative name servers for

each domain to keep track of their own changes, avoiding the need for a central register to be continually consulted and updated.
In general, the Domain Name System also stores other types of information, such as the list of mail servers that accept email for a given Internet domain. By providing a world-wide, distributed keyword-based redirection service, the Domain Name System is an essential component of the functionality of the Internet.

41.2 DNSv4/v6 Configuration Task List

To enable/disable DNS function
To configure/delete DNS server
To configure/delete domain name suffix
To delete the domain entry of specified address in dynamic cache
To enable DNS dynamic domain name resolution
Enable/disable DNS SERVER function
Configure the max number of client information in the switch queue
Configure the timeout value of caching the client information on the switch
Monitor and diagnosis of DNS function

1. To enable/disable DNS function

Command	Explanation
Global Mode	
ip domain-lookup **no ip domain-lookup**	To enable/disable DNS dynamic lookup function.

2. To configure/delete DNS server

Command	Explanation
Global Mode	
dns-server {<ip-address> | <ipv6-address>} [priority <value>] **no dns-server {<ip-address> | <ipv6-address>}**	To configure DNS server, the no form of this command deletes DNS server.

3. To configure/delete domain name suffix

Command	Explanation
Global Mode	
ip domain-list <*WORD*> **no ip domain-list <*WORD*>**	To configure/delete domain name suffix.

4. To delete the domain entry of specified address in dynamic cache

Command	Explanation
Admin Mode	
clear dynamic-host {<ip-address> | <ipv6-address> | all}	To delete the domain entry of specified address in dynamic cache.

5. To enable DNS dynamic domain name resolution

Command	Explanation
Global Mode	
dns lookup {ipv4 | ipv6} *<hostname>*	To enable DNS dynamic domain name resolution.

6. Enable/disable DNS SERVER function

Command	Explanation
Global Mode	
ip dns server **no ip dns server**	Enable/disable DNS SERVER function.

7. Configure the max number of client information in the switch queue

Command	Explanation
Global Mode	
ip dns server queue maximum <1-5000> **no ip dns server queue maximum**	Configure the max number of client information in the switch queue.

8. Configure the timeout value of caching the client information on the switch

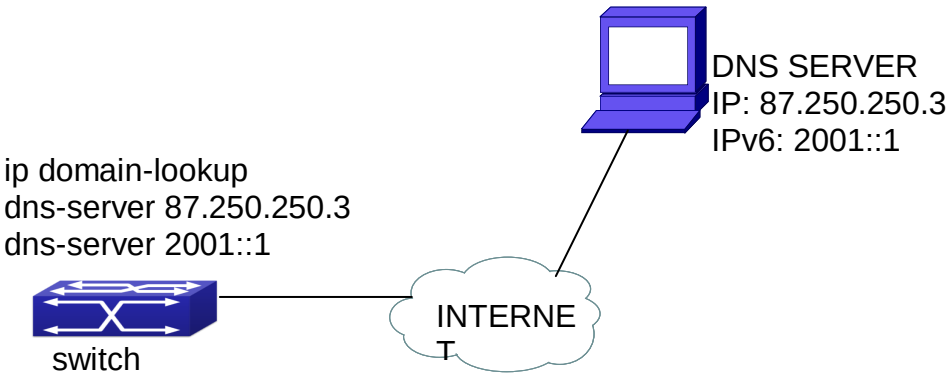
Command	Explanation
Global Mode	
ip dns server queue timeout <1-100> **no ip dns server queue timeout**	Configure the timeout value of caching the client information on the switch.

9. Monitor and diagnosis of DNS function

Command	Explanation
Admin Mode and Configuration Mode	
show dns name-server	To show the configured DNS server information.
show dns domain-list	To show the configured DNS domain name suffix information.
show dns hosts	To show the dynamic domain name information

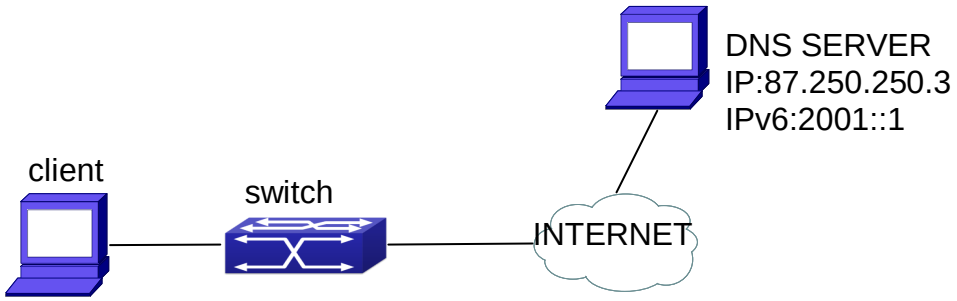
	of resolved by switch.
show dns config	Display the configured global DNS information on the switch.
show dns client	Display the DNS Client information maintained by the switch.
debug dns {all | packet [send | recv] | events | relay} **no debug dns {all | packet [send | recv] | events | relay}**	To enable/disable DEBUG of DNS function.

41.3 Typical Examples of DNS



DNS CLIENT typical environment

As shown in fig, the switch connected to DNS server through network, if the switch want to visit yandex website, it needn't to know the IPv4/IPv6 address of yandex website, only need is to record the domain name www.yandex.ru. The DNS server can resolute out the IPv4/IPv6 address of this domain name and send to switch, then the switch can visit www.yandex.ru correctly. The switch is configured as DNS client, basic configurations are as below: first to enable DNS dynamic domain name resolution function on switch, and configure DNS server address, then with some kinds of tools such as PING, the switch can get corresponding IPv4/IPv6 address with dynamic domain name resolution function.



The figure above is an application of DNS SERVER. Under some circumstances, the client PC doesn't know the real DNS SERVER, and points to the switch instead. The switch plays the role of a DNS SERVER in two steps: Enable the global DNS SERVER function, configure the IP address of the real DNS server. After the DNS SERVER function is globally enabled, the switch will look up its local cache when receiving a DNS request from a client PC. If there is a domain needed by the local client, it will directly answer the client's request; otherwise, the switch will relay the request to the real DNS server, pass the reply from the DNS Server to the client and record the domain and its IP address for a faster lookup in the future.

Switch configuration for DNS CLIENT:
Switch(config)# ip domain-lookup
Switch(config)# dns-server 87.250.250.3
Switch(config)# dns-server 2001::1
Switch#ping host www.yandex.ru
Switch#traceroute host www.yandex.ru
Switch#telnet host www.yandex.ru

Switch configuration for DNS SERVER:
Switch(config)# ip domain-lookup
Switch(config)# dns-server 87.250.250.3
Switch(config)# dns-server 2001::1
Switch(config)# ip dns server

41.4 DNS Troubleshooting

In configuring and using DNS, the DNS may fail due to reasons such as physical connection failure or wrong configurations. The user should ensure the following:
First make sure good condition of the TACACS+ server physical connection;
Second all interface and link protocols are in the UP state (use "**show interface**" command);
Then please make sure that the DNS dynamic lookup function is enabled (use the "ip domain-lookup" command) before enabling the DNS CLIENT function. To use DNS SERVER function, please enable it (use the "ip dns server" command);
Finally ensure configured DNS server address (use "**dns-server**" command), and the switch can ping DNS server;
If the DNS problems remain unsolved, please use debug DNS all and other debugging command and copy the DEBUG message within 3 minutes, send the recorded message to the technical service center of our company.

Chapter 42 Summer Time Configuration

42.1 Introduction to Summer Time

Summer time is also called daylight saving time, it is a time system for saving energy sources. In summer the time is advanced 1 hour to keep early hours, reduce the lighting, so as to save electrolighting. The rule that adopt summer time is different in each country. At present, almost 110 countries implement summer time.

Compare with the standard time, usually set summer time 1 hour late, for example, when summer time is implementing, 10:00 am of the standard time is considered 11:00 am of summer time.

42.2 Summer Time Configuration Task Sequence

1. Configure absolute or recurrent time range of summer time

Command	Explanation
Global Mode	
clock summer-time <*word*> absolute <*HH:MM*> <*YYYY.MM.DD*> <*HH:MM*> <*YYYY.MM.DD*> [<*offset*>] **no clock summer-time**	Set absolute time range of summer time, start and end summer time is configured with specified year.
clock summer-time <*word*> recurring <*HH:MM*> <*MM.DD*> <*HH:MM*> <*MM.DD*> [<*offset*>] **no clock summer-time**	Set recurrent time range of summer time, every year the summer time begins from the start time and end at the end time.
clock summer-time <*word*> recurring <*HH:MM*> <*week*> <*day*> <*month*> <*HH:MM*> <*week*> <*day*> <*month*> [<*offset*>] **no clock summer-time**	Set recurrent time range of summer time, every year the summer time begins from the start time and end at the end time.

42.3 Examples of Summer Time

Example1:

The configuration requirement in the following: The summer time from 23:00 on April 1th, 2012 to 00:00 on October 1th, 2012, clock offset as 1 hour, and summer time is named as 2012.

Configuration procedure is as follows:

Switch(config)# clock summer-time 2012 absolute 23:00 2012.4.1 00:00 2012.10.1

Example2:
The configuration requirement in the following: The summer time from 23:00 on the first Saturday of April to 00:00 on the last Sunday of October year after year, clock offset as 2 hours, and summer time is named as time_travel.

Configuration procedure is as follows:
Switch(config)#clock summer-time time_travel recurring 23:00 first sat apr 00:00 last sun oct 120

42.4 Summer Time Troubleshooting

If there is any problem happens when using summer time, please check whether the problem is caused by the following reasons:
Check whether command mode in global mode
Check whether system clock is correct

Chapter 43 Monitor and Debug

When the users configures the switch, they will need to verify whether the configurations are correct and the switch is operating as expected, and in network failure, the users will also need to diagnostic the problem. Switch provides various debug commands including ping, telnet, show and debug, etc. to help the users to check system configuration, operating status and locate problem causes.

43.1 Ping

Ping command is mainly used for sending ICMP query packet from the switches to remote devices, also for check the accessibility between the switch and the remote device. Refer to the Ping command chapter in the Command Manual for explanations of various parameters and options of the Ping command.

43.2 Ping6

Ping6 command is mainly used by the switch to send ICMPv6 query packet to the remote equipment, verifying the accessibility between the switch and the remote equipment. Options and explanations of the parameters of the Ping6 command please refer to Ping6 command chapter in the command manual.

43.3 Traceroute

Traceroute command is for testing the gateways through which the data packets travel from the source device to the destination device, so to check the network accessibility and locate the network failure.

Execution procedure of the Traceroute command consists of: first a data packet with TTL at 1 is sent to the destination address, if the first hop returns an ICMP error message to inform this packet can not be sent (due to TTL timeout), a data packet with TTL at 2 will be sent. Also the send hop may be a TTL timeout return, but the procedure will carries on till the data packet is sent to its destination. These procedures is for recording every source address which returned ICMP TTL timeout message, so to describe a path the IP data packets traveled to reach the destination.

Traceroute Options and explanations of the parameters of the Traceroute command please refer to traceroute command chapter in the command manual.

43.4 Traceroute6

The Traceroute6 function is used on testing the gateways passed through by the data packets from the source equipment to the destination equipment, to verify the accessibility and locate the network failure. The principle of the Traceroute6 under IPv6 is the same as that under IPv4, which adopts the hop limit field of the ICMPv6 and IPv6 header. First, Traceroute6 sends an IPv6 datagram (including source address, destination address and packet sent time) whose HOPLIMIT is set to 1. When first route on the path receives this datagram, it minus the HOPLIMIT by 1 and the HOPLIMIT is now 0. So the router will discard this datagram and returns with a “ICMPv6 time exceeded” message (including the source address of the IPv6 packet, all content in the IPv6 packet and the IPv6 address of the router). Upon receiving this message, the Traceroute6 sends another datagram of which the HOPLIMIT is increased to 2 so to discover the second router. Plus 1 to the HOPLIMIT every time to discover another router, the Traceroute6 repeat this action till certain datagram reaches the destination.
Traceroute6 Options and explanations of the parameters of the Traceroute6 command please refer to traceroute6 command chapter in the command manual.

43.5 Show

show command is used to display information about the system, port and protocol operation. This part introduces the **show** command that displays system information, other **show** commands will be discussed in other chapters.

Command	Explanation
Admin Mode	
show debugging	Display the debugging state.
show flash	Display the files and the sizes saved in the flash.
show history	Display the recent user input history command.
show history all-users [detail]	Show the recent command history of all users. Use **clear history all-users** command to clear the command history of all users saved by the system, the max history number can be set by **history all-users max-length** command.
show memory	Display content in specified memory area.
show running-config	Display the switch parameter configuration validating at current operation state.
show startup-config	Display the switch parameter configuration written in the Flash Memory at current operation state, which is normally the configuration file applied in next time the switch starts up.

show switchport interface [ethernet *<IFNAME>*]	Display the VLAN port mode and the belonging VLAN number of the switch as well as the Trunk port information.
show tcp **show tcp ipv6**	Display the TCP connection status established currently on the switch.
show udp **show udp ipv6**	Display the UDP connection status established currently on the switch.
show telnet login	Display the information of the Telnet client which currently establishes a Telnet connection with the switch.
show tech-support	Display the operation information and the state of each task running on the switch. It is used by the technicians to diagnose whether the switch operates properly.
show version	Display the version of the switch.
show temperature	Show CPU temperature of the switch.

43.6 Debug

All the protocols switch supports have their corresponding debug commands. The users can use the information from debug commands for troubleshooting. Debug commands for their corresponding protocols will be introduced in the later chapters.

43.7 System log

43.7.1 System Log Introduction

The system log takes all information output under it control, while making detailed catalogue, so to select the information effectively. Combining with Debug programs, it will provide a powerful support to the network administrator and developer in monitoring the network operation state and locating the network failures.

The switch system log has following characteristics

Log output from four directions (or log channels) of the Console, Telnet terminal and monitor, log buffer zone, and log host.

The log information is classified to four level of severities by which the information will be filtered

According to the severity level the log information can be auto outputted to corresponding log channel.

43.7.1.1 Log Output Channel

So far the system log can be outputted the log information through four channels:
Through Console port to the local console
Output the log information to remote Telnet terminal or monitor, this function is good for remote maintenance
Assign a proper log buffer zone inside the switch, for record the log information permanently or temporarily
Configure the log host, the log system will directly send the log information to the log host, and save it in files to be viewed at any time

Among above log channels, users rarely use the console monitor, but will commonly choose the Telnet terminal to monitor the system operation status. However information outputted from these channels are of low traffic capacity and can not be recorded for later view. The other two channels---the log buffer zone and log host channel are two important channels
SDRAM (Synchronous Dynamic Random Access Memory) and NVRAM (Non Vulnerable Random Access Memory) is provided inside the switch as two part of the log buffer zone, The two buffer zone record the log information in a circuit working pattern, namely when log information need to be recorded exceeds the buffer size, the oldest log information will be erased and replaced by the new log information, information saved in NVRAM will stay permanently while those in SDRAM will lost when the system restarts or encounter an power failure. Information in the log buffer zone is critical for monitoring the system operation and detecting abnormal states.

Note: the NVRAM log buffer may not exist on some switches, which only have the SDRAM log buffer zone.
It is recommended to use the system log server. By configuring the log host on the switch, the log can be sent to the log server for future examination.

43.7.1.2 Format and Severity of the Log Information

The log information format is compatible with the BSD syslog protocol, so we can record and analyze the log by the systlog (system log protect session) on the UNIX/LINUX, as well as syslog similar applications on PC.
The log information is classified into eight classes by severity or emergency procedure. One level per value and the higher the emergency level the log information has, the smaller its value will be. For example, the level of critical is 2, and warning is 4, debugging is leveled at 7, so the critical is higher than warnings which no doubt is high than debugging. The rule applied in filtering the log information by severity level is that: only the log information with level equal to or higher than the threshold will be outputted. So when the severity threshold is set to debugging, all information will be outputted and if set to critical, only critical, alerts and

emergencies will be outputted.

Follow table summarized the log information severity level and brief description. **Note:** these severity levels are in accordance with the standard UNIX/LINUX syslog.

Severity of the loginformation

Severity	Value	Description
emergencies	0	System is unusable
alerts	1	Action must be taken immediately
critical	2	Critical conditions
errors	3	Error conditions
warnings	4	Warning conditions
notifications	5	Normal but significant condition
informational	6	Informational messages
debugging	7	Debug-level messages

Right now the switch can generate information of following four levels

Restart the switch, mission abnormal, hot plug on the CHASSIS switch chips are classified critical

Up/down interface, topology change, aggregate port state change of the interface are notifications warnings

Outputted information from the CLI command is classified informational

Information from the debugging of CLI command is classified debugging

Log information can be automatically sent to corresponding channels with regard to respective severity levels. Amongst the debugging information can only be sent to the monitor. Those with the Informational level can only be sent to current monitor terminal, such as the information from the Telnet terminal configuration command can only be transmitted to the Telnet terminal. Warnings information can be sent to all terminal with also saved in the SDRAM log buffer zone. And the critical information can be save both in SDRAM and the NVRAM (if exists) besides sent to all terminals. To check the log save in SDRAM and the NVRAM, we can use the show logging buffered command. To clear the log save in NVRAM and SDRAM log buffer zone, we can use the clear logging command.

43.7.2 System Log Configuration

System Log Configuration Task Sequence:

1. Display and clear log buffer zone
2. Configure the log host output channel
3. Enable/disable the log executed-commands
4. Display the log source
5. Display executed-commands state

Display and clear log buffer zone

Command	Description
Admin Mode	
show logging buffered [level {critical | warnings} | range <*begin-index*> <*end-index*>]	Show detailed log information in the log buffer channel.
clear logging {sdram | nvram}	Clear log buffer zone information.

Configure the log host output channel

Command	Description
Global Mode	
logging {<ipv4-addr> | <ipv6-addr>} [facility <local-number>] [level <severity>] **no logging {<*ipv4-addr*> | <*ipv6-addr*>} [facility <*local-number*>]**	Enable the output channel of the log host. The “no” form of this command will disable the output at the output channel of the log host.
logging loghost sequence-number **no logging loghost sequence-number**	Add the loghost sequence-number for the log, the no command does not include the loghost sequence-number.

Enable/disable the log executed-commands

Command	Description
Global mode	
logging executed-commands {enable | disable}	Enable or disable the logging executed-commands

Display the log source

Command	Description
Admin and configuration mode	
show logging source mstp	Show the log information source of MSTP module.

Display executed-commands state

Command	Description
Admin mode	
show logging executed-commands state	Show the state of logging executed-

43.7.3 System Log Configuration Example

Example 1: When managing VLAN the IPv4 address of the switch is 100.100.100.1, and the IPv4 address of the remote log server is 100.100.100.5. It is required to send the log information with a severity equal to or higher than warnings to this log server and save in the log record equipment local1.

Configuration procedure:
Switch(config)#interface vlan 1
Switch(Config-if-Vlan1)# ip address 100.100.100.1 255.255.255.0
Switch(Config-if-Vlan1)#exit
Switch(config)#logging 100.100.100.5 facility local1 level warnings

Example 2: When managing VLAN the IPv6 address of the switch is 3ffe:506::1, and the IPv4 address of the remote log server is 3ffe:506::4. It is required to send the log information with a severity equal to or higher than critical to this log server and save the log in the record equipment local7.

Configuration procedure
Switch(config)#interface vlan 1
Switch(Config-if-Vlan1)#ipv6 address 3ffe:506::1/64
Switch(Config-if-Vlan1)#exit
Switch(config)#logging 3ffe:506::4 facility local7 level critical

Chapter 44 Reload Switch after Specified Time

44.1 Introduce to Reload Switch after Specifid Time

Reload switch after specified time is to reboot the switch without shutdown its power after a specified period of time, usually when updating the switch version. The switch can be rebooted after a period of time instead of immediately after its version being updated successfully.

44.2 Reload Switch after Specifid Time Task List

1. Reload switch after specified time

Command	Explanation
Admin mode	
reload after {[<*HH:MM:SS*>] [days <*days*>]}	Reload the switch after a specified time period.
reload cancel	Cancel the specified time period to reload the switch.

Chapter 45 Debugging and Diagnosis for Packets Received and Sent by CPU

45.1 Introduction to Debugging and Diagnosis for Packets Received and Sent by CPU

The following commands are used to debug and diagnose the packets received and sent by CPU, and are supposed to be used with the help of the technical support.

45.2 Debugging and Diagnosis for Packets Received and Sent by CPU Task List

Command	Explanation
Global Mode	
cpu-rx-ratelimit total <*packets*> **no cpu-rx-ratelimit total**	Set the total rate of the CPU receiving packets, the no command sets the total rate of the CPU receiving packets to default.
cpu-rx-ratelimit protocol <protocol-type> <packets> **no cpu-rx-ratelimit protocol [<*protocol- type*>]**	Set the max rate of the CPU receiving packets of the protocol type, the no command set the max rate to default.
clear cpu-rx-stat protocol [<*protocol-type*>]	Clear the statistics of the CPU received packets of the protocol type.
Admin Mode	
show cpu-rx protocol [<*protocol-type*>]	Show the information of the CPU received packets of the protocol type.
debug driver {receive|send} [interface {<*interface-name*> |all}] [protocol {<*protocol-type*> |discard |all}][detail]	Turn on the showing of the CPU receiving or sending packet informations.
no debug driver {receive |send}	Turn off the showing of the CPU receiving or sending packet informations.