CONFIGURATION GUIDE

Extreme SLX-OS IP Multicast Configuration Guide, 18s.1.01

Supporting the ExtremeSwitching SLX 9140 and SLX 9240 Switches

9035262
November 2018

Contents

Preface

Document conventions

The document conventions describe text formatting conventions, command syntax conventions, and important notice formats used in Extreme technical documentation.

Notes, cautions, and warnings

Notes, cautions, and warning statements may be used in this document. They are listed in the order of increasing severity of potential hazards.

NOTE
A Note provides a tip, guidance, or advice, emphasizes important information, or provides a reference to related information.

ATTENTION
An Attention statement indicates a stronger note, for example, to alert you when traffic might be interrupted or the device might reboot.

CAUTION
A Caution statement alerts you to situations that can be potentially hazardous to you or cause damage to hardware, firmware, software, or data.

DANGER
A Danger statement indicates conditions or situations that can be potentially lethal or extremely hazardous to you. Safety labels are also attached directly to products to warn of these conditions or situations.

Text formatting conventions

Text formatting conventions such as boldface, italic, or Courier font may be used to highlight specific words or phrases.

Format	Description
bold text	Identifies command names.
	Identifies keywords and operands.
	Identifies the names of GUI elements.
	Identifies text to enter in the GUI.
italic text	Identifies emphasis.
	Identifies variables.
	Identifies document titles.
`Courier font`	Identifies CLI output.

Format	Description
	Identifies command syntax examples.

Command syntax conventions

Bold and italic text identify command syntax components. Delimiters and operators define groupings of parameters and their logical relationships.

Convention	Description
bold text	Identifies command names, keywords, and command options.
italic text	Identifies a variable.
[ ]	Syntax components displayed within square brackets are optional. Default responses to system prompts are enclosed in square brackets.
{ **x** | **y** | **z** }	A choice of required parameters is enclosed in curly brackets separated by vertical bars. You must select one of the options.
x | **y**	A vertical bar separates mutually exclusive elements.
< >	Nonprinting characters, for example, passwords, are enclosed in angle brackets.
...	Repeat the previous element, for example, *member*[*member*...].
\	Indicates a "soft" line break in command examples. If a backslash separates two lines of a command input, enter the entire command at the prompt without the backslash.

Extreme resources

Visit the Extreme website to locate related documentation for your product and additional Extreme resources.

White papers, data sheets, and the most recent versions of Extreme software and hardware manuals are available at www.extremenetworks.com. Product documentation for all supported releases is available to registered users at www.extremenetworks.com/support/documentation.

Document feedback

Quality is our first concern at Extreme, and we have made every effort to ensure the accuracy and completeness of this document. However, if you find an error or an omission, or you think that a topic needs further development, we want to hear from you.

You can provide feedback in two ways:

- Use our short online feedback form at http://www.extremenetworks.com/documentation-feedback-pdf/
- Email us at internalinfodev@extremenetworks.com

Provide the publication title, part number, and as much detail as possible, including the topic heading and page number if applicable, as well as your suggestions for improvement.

Contacting Extreme Technical Support

As an Extreme customer, you can contact Extreme Technical Support using one of the following methods: 24x7 online or by telephone. OEM customers should contact their OEM/solution provider.

If you require assistance, contact Extreme Networks using one of the following methods:

- GTAC (Global Technical Assistance Center) for immediate support
  - Phone: 1-800-998-2408 (toll-free in U.S. and Canada) or +1 408-579-2826. For the support phone number in your country, visit: www.extremenetworks.com/support/contact.
  - Email: support@extremenetworks.com. To expedite your message, enter the product name or model number in the subject line.
- GTAC Knowledge - Get on-demand and tested resolutions from the GTAC Knowledgebase, or create a help case if you need more guidance.
- The Hub - A forum for Extreme customers to connect with one another, get questions answered, share ideas and feedback, and get problems solved. This community is monitored by Extreme Networks employees, but is not intended to replace specific guidance from GTAC.
- Support Portal - Manage cases, downloads, service contracts, product licensing, and training and certifications.

Before contacting Extreme Networks for technical support, have the following information ready:

- Your Extreme Networks service contract number and/or serial numbers for all involved Extreme Networks products
- A description of the failure
- A description of any action(s) already taken to resolve the problem
- A description of your network environment (such as layout, cable type, other relevant environmental information)
- Network load at the time of trouble (if known)
- The device history (for example, if you have returned the device before, or if this is a recurring problem)
- Any related RMA (Return Material Authorization) numbers

About This Document

Supported hardware and software

In those instances in which procedures or parts of procedures documented here apply to some devices but not to others, this guide identifies exactly which devices are supported and which are not.

Although many different software and hardware configurations are tested and supported by Extreme Networks, Inc. for this SLX-OS release, documenting all possible configurations and scenarios is beyond the scope of this document.

The following hardware platforms are supported by this release:

- ExtremeSwitching SLX 9140
- ExtremeSwitching SLX 9240

NOTE
Some of the commands in this document use a slot/port designation. Because the SLX 9140 and the SLX 9240 do not contain line cards, the slot designation must always be "0" (for example, 0/1 for port 1).

What's new in this document

The following table includes descriptions of new information added to this guide for SLX-OS Release 18s.1.01.

TABLE 1 Summary of enhancements in SLX-OS Release 18s.1.01

Feature	Description	Described in
Layer 2 multicast support over MCT	SLX-OS devices support Layer 2 multicast control packets over Multi-Chassis Trunk (MCT).	Layer 2 multicast support over MCT on page 18

IP Multicast

IP multicast overview

Multicast protocols allow a group or channel to be accessed over different networks by multiple stations (clients) for the receipt and transmission of multicast data. Distribution of stock quotes, video transmissions such as news services and remote classrooms, and video conferencing are all examples of applications that use multicast routing.

Extreme devices support the Internet Group Management Protocol (IGMP) and Multicast Listener Discovery (MLD).

The Internet Group Management Protocol (IGMP) is used by IP hosts to report their multicast group memberships to any immediately-neighboring multicast routers.

Multicast control packet and data forwarding through a Layer 2 switch is achieved by Layer 2 forwarding of the received multicast packets on all the member ports of the VLAN interfaces. This approach though simple is not bandwidth efficient, because only a subset of member ports may be connected to devices interested in receiving these multicast packets. In a worst-case scenario, the data gets forwarded to all port members of a VLAN with a large number of member ports even if only a single VLAN member is interested in receiving the data. Such scenarios can lead to loss of throughput for a switch upon receiving a high rate of multicast data traffic.

IGMP and MLD provide the functionality to save bandwidth and throughput by forwarding traffic to only interested receivers instead of all the member ports of the VLAN. IGMP snooping provides the specification for IPv4 and MLD snooping provides the specification for IPv6 data traffic forwarding.

MLD snooping is a multicast constraining mechanism that runs on Layer 2 or Layer 3 devices to manage and control IPv6 multicast groups. MLD snooping provides similar functionality for IPv6 as IGMP snooping for IPv4 by sending IPv6 multicast traffic only to the interested listeners. By listening to and analyzing MLD messages, a Layer 2 device running MLD snooping establishes mappings between ports and multicast MAC addresses or multicast IP addresses and forwards multicast data.

IP multicast message types

Multicast routers use IGMP or MLD to learn which groups have interested listeners on each of their attached physical networks. In any given subnet, one multicast router is elected to act as an IGMP or MLD querier.

The IGMP or MLD querier sends out the following types of queries to hosts:

- General query: Asks whether any host is listening to any group.
- Group-specific query: Asks whether any host is listening to a specific multicast group. This query is sent in response to a host leaving the multicast group and allows the router to quickly determine if any remaining hosts are interested in the group.

Hosts that are multicast listeners send the following kinds of messages:

- Membership report: Indicates that the host wants to join a particular multicast group.
- Leave report: Indicates that the host wants to leave a particular multicast group.

IPv4 Multicast Traffic Reduction

IGMP snooping overview

The forwarding of multicast control packets and data through a Layer 2 device configured with VLANs is most easily achieved by the Layer 2 forwarding of received multicast packets on all the member ports of the VLAN interfaces. However, this simple approach is not bandwidth efficient, because only a subset of member ports may be connected to devices interested in receiving those multicast packets. In a worst-case scenario, the data would get forwarded to all port members of a VLAN with a large number of member ports, even if only a single VLAN member is interested in receiving the data. Such scenarios can lead to loss of throughput for a device that gets hit by a high rate of multicast data traffic.

Internet Group Management Protocol (IGMP) snooping is a mechanism by which a Layer 2 device can effectively address this issue of inefficient multicast forwarding to VLAN port members. Snooping involves "learning" forwarding states for multicast data traffic on VLAN port members from the IGMP control (join/leave) packets received on them. The Layer 2 device also provides for a way to configure forwarding states statically through the CLI.

Multicast routing and IGMP snooping

Multicast routers use IGMP snooping to learn which groups have members on each of their attached physical networks. A multicast router keeps a list of multicast group memberships for each attached network, and a timer for each membership.

NOTE

"Multicast group memberships" means that at least one member of a multicast group on a given attached network is available.

There are two ways that hosts join multicast routing groups:

- By sending an unsolicited IGMP join request.
- By sending an IGMP join request as a response to a general query from a multicast router.

In response to the request, the device creates an entry in its Layer 2 forwarding table for that VLAN. When other hosts send join requests for the same multicast, the device adds them to the existing table entry. Only one entry is created per VLAN in the Layer 2 forwarding table for each multicast group.

VLANs can be configured as snooping only or routing with snooping. When Layer 3 multicast routing is enabled on a particular VE, snooping for the underlying VLAN is enabled implicitly. Explicit snooping can be enabled on a VLAN in addition to implicit snooping. Implicit snooping is by default IGMP snooping. With routing enabled on a VE, when explicit snooping is disabled, snooping reverts back to implicit snooping. This does not change the functionality in any way, but only removes the configuration. When routing is disabled on a VE where explicit snooping is configured, the routing side of the programming stops and the snooping side programming takes over. When routing is enabled, the Layer 3 IGMP querier takes precedence on that VLAN. When routing is disabled, and if the snooping querier is configured, then the snooping querier takes effect.

Enabling IGMP snooping

1. Enter the **configure terminal** command to access global configuration mode.

```
device# configure terminal
```

2. Enter VLAN configuration mode.

```
device(config)# vlan 1
device(config-vlan-1)
```

3. Enable IGMP snooping.

```
device(config-vlan-1)# ip igmp snooping enable
```

Configuring the IGMP snooping querier

If your multicast traffic is not routed because Protocol-Independent Multicast (PIM) is not configured, use the IGMP snooping querier in a VLAN.

The IGMP snooping querier sends out IGMP queries to trigger IGMP responses from devices that are to receive IP multicast traffic. The IGMP snooping querier listens for these responses to map the appropriate forwarding addresses.

Use the following procedure to configure the IGMP snooping querier.

1. Enter the **configure terminal** command to access global configuration mode.

```
device# configure terminal
```

2. Enter the **vlan** command with the VLAN number.

```
device(config)# vlan 25
```

3. Set the IGMP query interval for the VLAN.

```
device(config-Vlan-25)# ip igmp snooping query-interval 125
```

   The valid range is from 1 through 18000 seconds. The default is 125 seconds.

4. Set the last member query interval.

```
device(config-Vlan-25)# ip igmp snooping last-member-query-interval 1000
```

   The valid range is from 1000 through 25500 milliseconds. The default is 1000 milliseconds.

5. Set the last member query count.

```
device(config-Vlan-25)# ip igmp snooping last-member-query-count 3
```

   The valid range is from 2 through 10. The default is 2.

6. Set the startup query count.

```
device(config-Vlan-25)# ip igmp snooping startup-query-count 3
```

   The valid range is from 1 through 10. The default is 1.

7. Set the startup query interval.

```
device(config-Vlan-25)# ip igmp snooping startup-query-interval 200
```

   The valid range is from 1 through 450 seconds. The default is 1 second.

8. Set the Maximum Response Time.

```
device(config-Vlan-25)# ip igmp snooping query-max-response-time 10
```

   The valid range is from 1 through 25 seconds. The default is 10 seconds.

9. Configure the static Mrouter port.

```
device(config-Vlan-25)# ip igmp snooping mrouter interface ethernet 3/2
```

10. Configure a static IGMP group.

```
device(config-vlan-25)# ip igmp snooping static-group 225.0.0.1 interface ethernet 6/15
```

11. Set the snooping robustness variable.

```
device(config-Vlan-25)# ip igmp snooping robustness-variable 5
```

    The valid range is from 2 through 10. The default is 2.

12. You can stop the flooding of the unknown multicast traffic using the **ip igmp snooping restrict-unknown-multicast** command.

```
device(config-vlan-25)# ip igmp snooping restrict-unknown-multicast
```

13. Use the **ip igmp snooping fast-leave** command to enable fast leave processing.

```
device(config-vlan-25)# ip igmp snooping fast-leave
```

14. Activate the IGMP snooping querier functionality for the VLAN.

```
device(config-Vlan-25)# ip igmp snooping querier enable
```

    NOTE
    The IGMP snooping querier and the static mrouter can be configured together on a VLAN interface.

Monitoring IGMP snooping

Monitoring the performance of your IGMP traffic allows you to diagnose any potential issues on your device. This helps you utilize bandwidth more efficiently by setting the device to forward IP multicast traffic only to connected hosts that request multicast traffic.

Use the following commands to monitor IGMP snooping on the device; the commands do not need to be entered in any specific order.

1. Enter the **show ip igmp groups** command to display all information on IGMP multicast groups for the device. Use this command to display the IGMP database, including configured entries for all groups on all interfaces, all groups on specific interfaces, or specific groups on specific interfaces.

```
device# show ip igmp groups
Total Number of Groups: 2
IGMP Connected Group Membership
Group Address  Interface Uptime      Expires      Last Reporter   Version
225.1.1.1      vlan25    00:05:27     00:02:32     25.1.1.1202
Member Ports: eth 2/24
```

2. Enter the **show ip igmp snooping** command specifying the VLAN ID to view snooping configuration information such as snooping querier enable, snooping query interval, IGMP operation mode, PIM snooping configuration, and IGMP snooping configuration.

```
device# show ip igmp snooping

 Vlan ID: 10
 Multicast Router ports:  eth1/1
 Querier - Disabled
 IGMP Operation mode: IGMPv3
 Is Fast-Leave Enabled : Enabled
 Max Response time = 10
 Last Member Query Interval = 1
 Query interval = 125
 Number of Multicast Groups: 0
```

3. Enter the **show ip igmp statistics interface** command to display the IGMP statistics for a VLAN or interface.

```
device# show ip igmp statistics interface vlan 1

IGMP packet statistics for all interfaces in vlan 1:
IGMP Message type     Edge-Received   Edge-Sent   Edge-Rx-Errors   ISL Received
Membership Query                  0           0                0              0
V1 Membership Report              0           0                0              0
V2 Membership Report              0           0                0              0
Group Leave                       0           0                0              0
V3 Membership Report              0           0                0              0
PIM hello                         0           0                0              0

IGMP Error Statistics:
Unknown types            0
Bad Length               0
Bad Checksum             0
```

4. Enter the **show ip igmp interface** command to display the Layer 3 IGMP interface configuration information.

```
device# show ip igmp interface
Interface Ve100
IGMP enabled
IGMP query interval 30 seconds
IGMP other-querierinterval 65 seconds
IGMP query response time 10 seconds
IGMP last-member query interval 1 seconds
IGMP immediate-leave disabled
IGMP querier100.0.0.1(this system)
IGMP version 2
```

5. Enter the **show ip igmp snooping mrouter vlan** command to display mrouter port-related information.

```
device# show ip igmp snooping mrouter vlan 10
Vlan     Interface    Expires (Sec)
10       eth1/4       250
10       eth1/1       238
```

When you have reviewed the IGMP statistics for the device, refer to Enabling IGMP snooping on page 14 or Configuring the IGMP snooping querier on page 14 to make any needed corrections.

IPv4 Multicast Snooping

IGMP

The Internet Group Management Protocol (IGMP) allows an IPv4 system to communicate IP multicast group membership information to its neighboring routers. The routers, in turn, limit the multicast of IP packets with multicast destination addresses to only those interfaces on the router that are identified as IP multicast group members.

In IGMPv2, when a router sends a query to the interfaces, the clients on the interfaces respond with a membership report of multicast groups to the router. The router can then send traffic to these groups, regardless of the traffic source. When an interface no longer needs to receive traffic from a group, it sends a leave message to the router, which in turn sends a group-specific query to that interface to see if any other clients on the same interface are still active.

There are different types of query messages:

- A "General Query" is sent by a multicast router to learn the complete multicast reception state of the neighboring interfaces. In a General Query, both the Group Address field and the Number of Sources (N) field are zero.
- A "Group-Specific Query" is sent by a multicast router to learn the reception state, with respect to a *single* multicast address, of the neighboring interfaces. In a Group-Specific Query, the Group Address field contains the multicast address of interest, and the Number of Sources (N) field contains zero.

Default IGMP version

IGMPv2 is enabled by default only when snooping or multicast routing are enabled on the system.

Also, you can specify what version of IGMP you want to run on a device on a per-VLAN basis. If you do not specify an IGMP version, IGMPv2 is used.

Compatibility with IGMPv1 and IGMPv2

Different multicast groups, interfaces, and routers can run their own versions of IGMP. The version of IGMP is reflected in the membership reports that the hosts send to the router. Routers and interfaces must be configured to recognize the version of IGMP you want them to process.

An interface or router sends the queries and reports that include its IGMP version specified on it. The interface may recognize a query or report that has a different version. For example, an interface running IGMPv2 can recognize IGMPv3 packets, but cannot process them. When the router sends out IGMP queries over an IGMPv2 interface, the equal or lower version of reports is supported, but a higher version of reports is not supported.

The version of IGMP can be specified per interface (physical port or virtual routing interface), and per physical port within a virtual routing interface.

Layer 2 multicast support over MCT

SLX-OS devices support Layer 2 multicast control packets over Multi-Chassis Trunk (MCT).

Multicast state information is synced between the MCT peers using MP-BGP EVPN transport. Multicast protocol packets are not sent on the peer link unless required. MCT does not support Muilticast Listener Discovery.

IGMP protocol packets are of two types:

1. Membership query
2. Membership reports - In a report, the multicast address field is that of the specific multicast address to which the sender is listening.

IGMP query packet processing on MCT

For IGMP queries, each EVI has BUM suppressed MGID associated, IGMP query packets need to be transmitted on ICL to address the following scenarios:

- The querier connected to only one of the MCT peer switches becomes the elected querier.
- Only one of the peer switches is configured as a querier.
- The switch ages out IGMP routes if memberships are not confirmed within the timeout interval. Although query packets are received on the MCT peer link, the mrouter port is not learned or considered on that peer link.

IGMP membership reports

For IGMP reports and leaves, note the following:

- Traditionally, each peer switch learns about Layer 2 multicast memberships by snooping the IGMP/MLD membership reports. The membership reports are then flooded on multicast router (mrouter) ports.
- For MCT, because an mrouter port is not learned on the peer link, membership reports are not flooded between the peer switches. Peer switches exchange learned routes by using EVPN NLRI messages between BGP peers running on the MCT cluster control VLAN.
- BGP on the peer switch communicates with the IGMP routes to the multicast module to add/delete group memberships, and allocates or rejects the MGIDs and installs or uninstalls the routes in hardware.
- If a general query or group-specific query is received from any port other than a peer link, each peer switch generates a proxy report for the IGMP routes learned from EVPN NLRI.

Duplicate IGMP query packets on CCEP

If an IGMP query is configured on both MCT peers of a member VLAN, duplicate IGMP query packets reach clients connected to the MCT domain by means of CCEP/CEP.

Multicast modules use inclusive multicast routes of EVPN NLRI for the member VLAN to exchange the multicast IGMP querier configuration on the member VLANs. The MCT peer that is the DF of the member VLAN is elected as the IGMP querier.

Multicast modules handle IGMP querier election for the member VLAN in peer link up/down scenarios.

Leave membership report

When fast-leave is not configured and an MCT peer receives a leave membership report from one of its clients for group G, the switch/router informs other MCT peers about the group-specific query and latency by using IGMP leave sync route. The peer switch, which runs the IGMP querier, sends group-specific queries and group queries.

Mrouter synchronization

Mrouter synchronization helps in achieving optimal path selection for unknown multicast traffic and optimal MP-BGP message exchange between MCT peers.

BGP handling of EVPN IGMP routes

In DC applications, EVPN is used as way to standardize inter-POD communication for both intra-DC and inter-DC applications.

A subnet can span across multiple PODs and DCs. EVPN provides a robust multitenant solution with extensive multihoming capabilities to stretch a subnet (VLAN) across multiple PODs and DCs. There can be many hosts/VMs (up to several hundreds) attached to a subnet that is stretched across several PODs and DCs. These hosts/virtual machines express their interests in multicast groups on a given subnet/VLAN by sending IGMP membership reports (Joins) for their interested multicast group(s). Also, an IGMP router (for example, IGMPv1) periodically sends membership queries to find out whether there are hosts on that subnet that are still interested in receiving multicast traffic for that group. As with the ARP/ND suppression mechanism in EVPN to reduce the flooding of ARP messages over EVPN, it is also desired to have a mechanism to reduce the flood of IGMP messages (both Queries and Reports) in EVPN. This is achieved through IGMP Join Sync and IGMP Leave Sync EVPN routes as specified in the draft *draft-ietf-bess-evpn-igmp-mld-proxy-00*.

TABLE 2 EVPN Routes

EVPN route type	Route type name	Description
7	IGMP Join Sync Route	To exchange (S,G)/(*,G) learned on BGP EVPN peers.
8	IGMP Leave Sync Route	To exchange group leave between BGP EVPN peers.

IGMP join sync route

IGMP allows a network host to inform a router that it is interested in receiving a particular multicast stream.

To begin, the multicast group is assigned a multicast address (that is, an IP address in the 224.0.0.0/4 class D address space). Hosts registered to receive the stream join the group by sending an IGMP Report to the upstream multicast router. The router then adds that group to the list of multicast groups that should be forwarded onto the local subnet.

The router does not maintain state about which hosts on the subnet are to receive traffic for the group. Instead, the router continues to send traffic to the subnet until either a timeout value expires or there are no more hosts in that group on the subnet. IGMP Join Sync route packets are described in the following table.

TABLE 3 IGMP Join Sync route packets

Packets	Description
RD (8 Octets)	Route Distinguisher
Ethernet Segment Identifier (10 Octets)	0 if (S,G) / (*,G) is learned on CEP
Ethernet Tag ID (4 octets)	0 or optionally, set to the VLAN ID over which a multicast route is learned
Multicast Source Length (1 octet)	32 for IPv4 address, 128 for IPv6 address and 0 for wildcard address (*)

TABLE 3 IGMP Join Sync route packets (continued)

Packets	Description
Multicast Source Address (variable)	IP address of multicast source
Multicast Group Length (1 octet)	32 for IPv4 address, 128 for IPv6 address
Multicast Group Address (variable)	IP address of multicast group
Originator Router Length (1 octet)	32 for IPv4 address, 128 for IPv6 address
Originator Router Address (variable)	The IP address of the originating router
Flags (1 octets) (optional)	The flag fields are defined in the following table.

Parameters and values are as follows:

- Querier Config Sync: Multicast Group is set to 224.0.0.2
- Mrouter Sync: Multicast Group is set to 224.0.0.1
- Originator router length: 32 for IPv4 address and 128 for IPv6 address
- Originator router address: IP address of the originating router

The following table lists the flags and associated values.

TABLE 4 Flags and values

0	1	2	3	4	5	6	7
Reserved	Querier config synch	Mrouter synch	PIM snooping	IE	V3	V2	V1
	Bit 1 indicates QuerierConfigSynch.	Bit 2 indicates MrouterSync.	Bit 3 indicates PIM Snooping.	Bit 4 indicates whether the (S, G) information carries within the route-type of Include Group type (bit value 0) or an Exclude Group type (bit value 1). The Exclude Group type bit should be ignored if bit 5 is set in case of IGMP Version 2 & IGMP Version 1 & MLD Version 1.	Bit 5 indicates support for IGMP/MLD version 3.	Bit 6 indicates support for IGMP/MLD version 2.	Bit 7 indicates support for IGMP/MLD version 1.

Note the following bit values:

- Bits 1, 2, 3 are proprietary fields.
- Bit 3 indicates PIM Snooping.
- Bit 2 indicates MrouterSync.
- Bit 1 indicates QuerierConfigSync.

IGMP leave sync route

When a host no longer wants to receive multicast traffic, it sends the router an IGMP Leave message.

After receiving this message, the router sends a query to the local subnet to determine whether any group members remain, sending the message to all hosts on the subnet at the multicast All-Hosts address (224.0.0.1). If any host responds, the router continues to send to the group; if not, the router removes the multicast group from its forwarding list and stops sending to the group. The following table describes the IGMP Leave Sync route packets.

TABLE 5 IGMP Leave Sync route packets

Packets	Description
RD (8 Octets)	
Ethernet Segment Identifier (10 octets)	0 if (S,G) / (*,G) is learned on CEP
Ethernet Tag ID (4 octets)	0 or optionally, set to the VLAN ID over which the multicast route is learned
Multicast Source Length (1 octet)	32 for IPv4 address, 128 for IPv6 address and 0 for wildcard address (*)
Multicast Source Address (variable)	IP address of multicast source
Multicast Group Length (1 octet)	32 for IPv4 address, 128 for IPv6 address
Multicast Group Address (variable)	IP address of multicast group
Originator Router Length (1 octet)	32 for IPv4 address, 128 for IPv6 address
Originator Router Address (variable)	The IP address of the originating router
Flags (1 octets) (optional)	The flag fields are defined in the following table

TABLE 6 Flags

0	1	2	3	4	5	6	7
Reserved	Reserved	Reserved	Reserved	IE	V3	V2	V1
				Bit 4 indicates whether the (S, G) information carries within the route-type of Include Group type (bit value 0) or an Exclude Group type (bit value 1). The Exclude Group type bit should be ignored if bit 5 is set in case of IGMP Version 2 & IGMP Version 1 & MLD Version 1.	Bit 5 indicates support for IGMP/MLD version 3.	Bit 6 indicates support for IGMP/MLD version 2.	Bit 7 indicates support for IGMP/MLD version 1.

IMPORTANT

Bits 0 to 3 are reserved and used.

NOTE

Leave Group synchronization and maximum response times are used during Leave Group synch procedures.

IGMP join and leave procedure

IGMP routes are originated by MLD and are sent to BGP through RibLib to be transported to the BGP EVPN peers.

These routes are carried in BGP EVPN NLRI as Type 7 and Type 8 routes. BGP adds the EVI-RT extended community to the EVPN NLRI and transports the route to the EVPN peers. An EVPN configuration must be present for the specified EVI.

EVPN IGMP routes received from the remote peers are validated against import rules and added to VPN table in BGP only if the validation passes. The routes are later imported into the MAC VRF table, if the RT in the route matches the RTs configured for a given EVI (VLAN/BD). Consolidation of the routes from different sources (RDs) takes place in BGP after the route passes the route target and import filtering checks.

The routes are then installed in BGP and are also forwarded to MLD through RibLib. Installed routes are also forwarded to other EVPN peers. BGP EVPN should be configured to support the IGMP routes.

NOTE
No new configuration is needed in BGP to support the IGMP routes.

EVI Route Target extended community

The EVI Route Target (RT) is a new EVPN extended community of Type 6 and a subtype yet to be defined by IANA. However, EVI-RT extended community is NOT supported for IGMP routes.

Instead, the Route Target extended community with Type 0x00 and Subtype 0x02 is supported. This extended community carries the RT associated with the EVI (VLAN/BD), so that the receiving PE can identify the EVI properly.

ES-Import Route Target extended community

ES-Import (ES-I) Route Target (RT) is another EVPN extended community of Type 0x06 and Subtype 0x02. The 6-byte value calculated is based on the ES-Import.

The IGMP Join and Leave Synch routes carry the ES-Import RT for the ES on which the IGMP membership report was received. Thus, it may go only to the PEs attached to that ES (and not to any other PEs).

Encap Type support

Only NSH tunnel encapsulation type is supported. MPLS and VXLAN types are not supported.

Traffic forwarding path for Layer 2 multicast

Both peers are updated with (*,G) membership for CCEP.

When a receiver connected on CCEP sends a membership report to join group G, the designated forwarder (DF) election of the CCEP and MGID update by multicast module prevents duplication of multicast data packets destined to group G on CCEP as well as on the peer link.

DF election is always honored for programming receivers on CCEP in MGID. However, the path given by DF election may not be optimal, as it might direct the multicast data traffic originating at one peer switch to a receiver on CCEP over the peer link, even though (*,G) membership does not include CEP on the peer switch that does not host the source.

Optimal traffic forwarding

When a group is learned over member VLANs, the DF for the IGMP/MLD route is elected by hashing on the IVID of the member VLAN, Source-IP, and Group-IP.

The OIF list of IGMP/MLD route on DF includes the following:

- Receivers connected through CCEP
- ICL, if its MCT peer has receivers connected through CEP
- Receivers connected through local CEP

The OIF list of IGMP/MLD routes on the non-DF includes the following:

- ICL to redirect the multicast stream to the DF of the stream
- Receivers connected through CEP

Data encapsulation of Layer 2 multicast traffic on ICL

Data encapsulation of Layer 2 multicast from CEP/CCEP received on a member VLAN is similar to that for Layer 2 flooding traffic.

- If the CCP is down, it forwards locally.
- If the remote CCEP is down, it forwards locally.
- If the local CCEP is down, it does not forward locally.
- If the ingress is the CEP, it forwards locally.
- If the ingress is the ICL, it does not forward locally.
- If the ingress is a different CCEP, it forwards locally.

Displaying Layer 2 over MCT status and clearing statistics

This section presents useful **show** and **clear** commands for this feature.

Show commands

The following examples present a variety of useful **show** displays.

show ip igmp groups cluster

```
device# show ip igmp groups cluster 1

Total Number of IGMP routes: 500
IGMP Connected Group Membership
Group Address    Source Address    Interface    Member    Type    Filter Mode    Mcast DF    Peer Address
------------------------------------------------------------------------------------------------------
227.0.0.1        Nil               Vlan 150     ICL        CEP     EXCLUDE        -
2.2.2.10
227.0.0.2        Nil               Vlan 150     ICL        CEP     EXCLUDE        -
2.2.2.10
227.0.0.3        Nil               Vlan 150     ICL        CEP     EXCLUDE        -
2.2.2.10
228.0.0.1        Nil               Vlan 150     Po 6       CCEP    EXCLUDE        DF          2.2.2.10
228.0.0.2        Nil               Vlan 150     Po 6       CCEP    EXCLUDE        DF          2.2.2.10
228.0.0.3        Nil               Vlan 150     Po 6       CCEP    EXCLUDE        Non-DF      2.2.2.10
228.0.0.5        Nil               Vlan 150     Po 6       CCEP    EXCLUDE        Non-DF      2.2.2.10
228.0.1.6        Nil               Vlan 150     Po 6       CCEP    EXCLUDE        Non-DF      2.2.2.10
```

show bgp evpn routes type igmp-join-sync

```
device# show bgp evpn routes type igmp-join-sync
  Total number of BGP EVPN Igmp Join Sync Routes : 2
Status A:AGGREGATE B:BEST b:NOT-INSTALLED-BEST C:CONFED_EBGP D:DAMPED
       E:EBGP H:HISTORY I:IBGP L:LOCAL M:MULTIPATH m:NOT-INSTALLED-MULTIPATH
       S:SUPPRESSED F:FILTERED s:STALE
       Prefix             Next Hop        MED        LocPrf     Weight Status
Route Distinguisher: 19.1.2.3:32868
1      IGMPJoinSyncPrefix4:[0](100.1.2.3,234.1.2.3):19.1.2.3 (esi 00.010203040506070809)
                          19.1.2.3        0          100        0      BI
         AS_PATH:
2      IGMPJoinSyncPrefix6:[0](2001::4,ff03::1):19.1.2.3 (esi 02.010203040506070809)
                          19.1.2.3        0          100        0      BI
         AS_PATH:
```

show bgp evpn routes type igmp-join-sync detail

```
device# show bgp evpn routes type igmp-join-sync detail
Total number of BGP EVPN Igmp Join Sync Routes :
2
Status A:AGGREGATE B:BEST b:NOT-INSTALLED-BEST C:CONFED_EBGP
D:DAMPED
```

```
       E:EBGP H:HISTORY I:IBGP L:LOCAL M:MULTIPATH m:NOT-INSTALLED-
MULTIPATH
       S:SUPPRESSED F:FILTERED
s:STALE
Route Distinguisher:
19.1.2.3:32868
1       Prefix: IGMPJoinSyncPrefix4:[0](100.1.2.3,234.1.2.3):19.1.2.3 (esi 00.010203040506070809),  Status:
BI,  Age: 0h4m45s
         NEXT_HOP: 19.1.2.3, Metric: 1, Learned from Peer: 19.1.2.3
(100)
          LOCAL_PREF: 100,  MED: 0,  ORIGIN: incomplete,  Weight:
0

AS_PATH:

            Extended Community: RT :1073741924 ExtCom:03:0c:
00:00:00:00:00:0a
            Extended Community: ExtCom: Tunnel Encapsulation (Type
MPLS)
            RD:
19.1.2.3:32868

2       Prefix: IGMPJoinSyncPrefix6:[0](2001::4,ff03::1):19.1.2.3 (esi 02.010203040506070809),  Status:
BI,  Age: 0h4m30s
         NEXT_HOP: 19.1.2.3, Metric: 1, Learned from Peer: 19.1.2.3
(100)
          LOCAL_PREF: 100,  MED: 0,  ORIGIN: incomplete,  Weight:
0

AS_PATH:

            Extended Community: RT :1073741924 ExtCom:03:0c:
00:00:00:00:00:0a
            Extended Community: ExtCom: Tunnel Encapsulation (Type
MPLS)
            RD:
19.1.2.3:32868
```

show bgp evpn routes type igmp-join-sync brief

```
device# show bgp evpn routes type igmp-join-sync brief
Total number of BGP EVPN Igmp Join Sync Routes : 2
Status codes: s suppressed, d damped, h history, * valid, > best, i internal, S stale
Origin codes: i - IGP, e - EGP, ? - incomplete
    Network            Next Hop        MED        LocPrf     Weight Path
Route Distinguisher: 19.1.2.3:32868
*>i IGMPJoinSyncPrefix4:[0](100.1.2.3,234.1.2.3):19.1.2.3 (esi 00.010203040506070809)
                          19.1.2.3        0          100        0      ?
*>i IGMPJoinSyncPrefix6:[0](2001::4,ff03::1):19.1.2.3 (esi 02.010203040506070809)
                          19.1.2.3        0          100        0      ?
```

Clear commands

The following table lists useful **clear** commands.

TABLE 7 Show and clear commands

Command	Description
clear ip igmp groups cluster	Clears statistics for all IPv4 groups learned through MCT on a cluster, or for groups learned on the CCEP.
clear ipv6 mld groups cluster	Clears statistics for all IPv6 groups learned through MCT on a cluster, or for groups learned on the CCEP.
clear bgp evpn l2routes type igmp-join-sync	Adds support for IGMP Join Sync.
clear bgp evpn l2routes type igmp-leave-sync	Adds support for IGMP Leave Sync.
clear bgp evpn routes type igmp-join-sync	Adds support for IGMP Join Sync.

TABLE 7 Show and clear commands (continued)

Command	Description
clear bgp evpn routes type igmp-leave-sync	Adds support for IGMP Leave Sync.
clear bgp evpn local routes type igmp-join-sync	Adds support for IGMP Join Sync.
clear bgp evpn local routes type igmp-leave-sync	Adds support for IGMP Leave Sync.

IPv6 Multicast VLAN Traffic Reduction

MLD snooping overview

Multicast Listener Discovery (MLD) snooping is a multicast-constraining mechanism that runs on Layer 2 or Layer 3 devices to manage and control IPv6 multicast groups.

A Layer 2 switch forwards all multicast control packets and data received on all the member ports of a VLAN interface. This approach, though simple, is not bandwidth efficient, because only a subset of member ports may be connected to devices interested in receiving those multicast packets. In the worst-case scenario the data are forwarded to all port members of a VLAN with a large number of member ports, even if only a single VLAN member is interested in receiving the data. Such scenarios can lead to loss of throughput for a switch when it receives high-rate multicast data traffic.

MLD snooping provides functionality for IPv6 that is similar to IGMP snooping for IPv4, by sending IPv6 multicast traffic only to interested listeners. By listening to and analyzing MLD messages, a Layer 2 device running MLD snooping establishes mappings between ports and multicast MAC addresses or multicast IP addresses, and forwards multicast data accordingly. Multicast routers in a network are found by means of either static configuration, dynamic learning, or PIM hello-based mrouter detection.

> NOTE
> This release supports the IPv6 version of MLDv1 snooping.

In any given subnet, one multicast router is elected to act as an MLD querier. The MLD querier sends out the following types of queries to hosts:

- General query: Querier asks whether any host is listening to any group.
- Group-specific query: Querier asks whether any host is listening to a specific multicast group. This query is sent in response to a host leaving the multicast group and allows the router to determine quickly whether any remaining hosts are interested in the group.

Hosts that are multicast listeners send the following kinds of messages:

- Report message: Indicates that the host wants to join a particular multicast group.
- Done message: Indicates that the host wants to leave a particular multicast group.

MLD traffic is forwarded as follows:

- MLD general queries received on a multicast-router interface are forwarded to all other interfaces in the VLAN.

- MLD group-specific queries received on a multicast-router interface are forwarded to only those interfaces in the VLAN that are members of the group.
- MLD report or done messages received on a host interface are forwarded to multicast-router interfaces in the same VLAN, but not to other host interfaces in the VLAN.
- Proxy MLD membership reports received with a null source IP address are accepted, to support report suppression.
- All unrecognized MLD packets are flooded to all (STP) unblocked member ports of the VLAN, to ensure that no data traffic is black-holed.

Data forwarding rules ensure that the multicast traffic received at the switch is forwarded to all interested downstream port members. Forwarding rules can be based on either the Layer 3 multicast destination IP group address or the Layer 2 destination MAC address.

- If a switch is already in a learned multicast group, multicast packets are forwarded only to those host interfaces in the VLAN that are members of the multicast group and to all multicast-router interfaces in the VLAN.
- If a switch is not in a learned multicast group, multicast packets for a group that has no current members are flooded to all member ports of the VLAN, as well as to all multicast-router interfaces in the VLAN.

The remainder of this section presents the tasks related to MLD configuration that are supported in this release.

Enabling and disabling MLD snooping at the VLAN level

You can enable or disable MLD snooping at the VLAN level.

1. Enter the configure terminal command to access global configuration mode.

```
device# configure terminal
```

2. Enter the VLAN configuration mode.

```
device(config)# vlan 1
device(config-vlan-1)
```

3. Enter the **ipv6 mld snooping enable** command.

```
device(config-vlan-1# ipv6 mld snooping enable
```

4. Enter the **no** form of the command to disable MLD snooping.

```
device(config-vlan-1# no ipv6 mld snooping enable
```

Enabling and disabling MLD querier functionality on a VLAN

You can use the MLD querier functionality to support MLD snooping on a VLAN where PIM and MLD are not enabled (for example, because multicast traffic does not need to be routed). MLD querier functionality is disabled by default.

To enable this functionality, use the **ipv6 mld snooping querier enable** command on a VLAN interface:

```
device(config-Vlan-2000)# ipv6 mld snooping querier enable
```

To disable this functionality, use the **no ipv6 mld snooping querier enable** command on a VLAN interface:

```
device(config-Vlan-2000)# no ipv6 mld snooping querier enable
```

Configuring and unconfiguring an MLD static group on a VLAN

You can forward traffic statically for a multicast group onto a specified interface, so that the interface behaves as if MLD were enabled.

1. To enable this functionality, use the **ipv6 mld snooping static-group** command on a VLAN interface, then select a multicast address to be joined, as well as a physical interface, as in the following example:

   ```
   device(config-Vlan-2000)# ipv6 mld snooping static-group ff1e::1 interface ethernet 0/1
   ```

2. To disable this functionality, use the **no ipv6 mld snooping static-group** command on a VLAN interface, as in the following example:

   ```
   device(config-Vlan-2000)# no ipv6 mld snooping static-group ff1e::1 interface ethernet 0/1
   ```

Enabling and disabling MLD fast-leave on a VLAN

MLD fast-leave allows a group entry to be removed immediately from the receiver as soon as a done message is received, as long as the receiver is the only one on the segment that is subscribed to a group. This minimizes the leave latency of group memberships on an interface, because the device does not send group-specific queries. As a result, the group entry is removed from the multicast forwarding table as soon as a group done (leave) message is received.

> NOTE
> Use the **ipv6 mld snooping fast-leave** command only if there is one receiver behind the interface for a given group.

Use the **ipv6 mld snooping fast-leave** command on a VLAN interface to enable MLD fast-leave.

```
device(config-Vlan-2000)# ipv6 mld snooping fast-leave
```

Use the **no ipv6 mld snooping fast-leave** command on a VLAN interface to disable MLD fast-leave.

```
device(config-Vlan-2000)# no ipv6 mld snooping fast-leave
```

Configuring the MLD query interval

You can configure the frequency at which MLD host query messages are sent. Larger values cause queries to be sent less often.

To set the MLD query interval, use the **ipv6 mld snooping query-interval** command on a VLAN interface, as in the following example:

```
device(config-Vlan-2000)# ipv6 mld snooping query-interval 1200
```

> NOTE
> The value set by the **ipv6 mld snooping query-interval** command must be greater than the query maximum response time, set by the **ipv6 mld query-max-response-time** command. Refer to the *Extreme SLX-OS Command Reference* for Extreme SLX 9140 and 9240 switches for all ranges and defaults for the commands in this section.

To restore the MLD query interval default value, use the **no ipv6 mld snooping query-interval** command on a VLAN interface:

```
device(config-Vlan-2000)# no ipv6 mld snooping query-interval
```

Configuring the MLD last-member query interval

You can set the frequency at which MLD last-member query messages are sent. This is the interval for the response to a query sent after a host leave message is received from the last known active host on the subnet. The group is deleted if no reports are received in this interval. This interval adjusts the speed at which messages are transmited on the subnet. Smaller values detect the loss of a group member faster.

> NOTE
> If the last-member query interval is not configured explicitly, the value is taken from the robustness variable.

To set the MLD last-member query interval, use the **ipv6 mld snooping last-member-query-interval** command on a VLAN interface:

```
switch(config-Vlan-2000)# ipv6 mld snooping last-member-query-interval 1500
```

To restore the default value, use the **no ipv6 mld snooping last-member-query-interval** command on a VLAN interface:

```
switch(config-Vlan-2000)# no ipv6 mld snooping last-member-query-interval
```

Configuring the MLD last-member query count

You can set the number of times that an MLD query is sent in response to a host leave message. This is the number of times, separated by the last-member query-response interval (configured by the **ipv6 mld last-member-query-interval** command), that an MLD query is sent in response to a host leave message from the last known active host on the subnet.

> NOTE
> If this interval is not configured explicitly, the value is taken from the robustness variable.

To change the MLD last-member query count from the default, use the **ipv6 mld last-member query count** command on a VLAN:

```
device(config-Vlan-2000)# ipv6 mld snooping last-member-query-count 3
```

To restore the default value, use the **no ipv6 mld last-member-query-count** command on a VLAN:

```
device(config-Vlan-2000)# no ipv6 mld snooping last-member-query-count
```

Configuring the MLD query maximum response time

You can configure the maximum response time for IPv6 MLDv1 snooping MLD queries for a specific VLAN interface:

```
device(config)# vlan 2
device(config Vlan-2)# ipv6 mld snooping query-max-response-time 15
```

> NOTE
> Larger values spread out host responses over a longer time. The value set by this command must be less than the general query interval, set by the **ipv6 mld query-interval** command.

To restore the default value, use the **no ipv6 mld query-max-response-time** command on a VLAN interface:

```
device(config-Vlan-2)# no ipv6 mld query-max-response-time
```

Configuring the MLD snooping robustness variable

A robustness value can be configured to compensate for packet loss in congested networks. This value determines the number of general MLD snooping queries that are sent before a multicast address is aged out for lack of a response. The default is 2.

To change the default robustness variable on a VLAN, use the **ipv6 mld snooping robustness-variable** command, as in the following example:

```
switch(config-Vlan-2000)# ipv6 mld snooping robustness-variable 7
```

To restore the default value, use the **no ipv6 snooping mld robustness-variable** command on a VLAN interface, as in the following example:

```
switch(config-Vlan-2000)# no ipv6 snooping mld robustness-variable
```

Configuring the MLD startup query count

The IPv6 MLDv1 startup query count is the number of queries that are separated by the startup interval. The default is 1.

Do the following to change the startup-query interval on a VLAN, as in the following example.

```
device(config-Vlan-2000)# ipv6 mld snooping startup-query-count 2
```

To restore the default value, use the **no ipv6 mld snooping startup-query-count** command on a VLAN interface, as in the following example:

```
device(config-Vlan-2000)# no ipv6 mld snooping startup-query-count
```

Configuring the MLD startup query interval

You can change the query interval between the general queries that are sent by the querier on startup. The default interval is 1. The querier may be the MLD snooping querier or an external querier.

To change the startup-query interval on a VLAN use **ipv6 mld startup-query-interval** command.

```
device(config-Vlan-2000)# ipv6 mld snooping startup-query-interval 2
```

To restore the default value, use the **no ipv6 mld snooping startup-query-interval** command on a VLAN interface:

```
device(config-Vlan-2000)# no ipv6 mld snooping startup-query-interval
```

Configuring a VLAN port member to be a multicast router port

You can configure a VLAN port member to be a multicast router (mrouter) port.

To configure a VLAN port member to be a multicast router (mrouter) port., use the **ipv6 mld snooping mrouter interface** command on a VLAN interface:

```
device(config-Vlan-2000)# ipv6 mld snooping mrouter interface ethernet 0/1
```

To disable the VLAN port member from being an mrouter port., use the **no ipv6 mld snooping mrouter interface** command on a VLAN interface:

```
device(config-Vlan-2000)# no ipv6 mld snooping mrouter interface ethernet 0/1
```

Monitoring and managing MLD snooping

You can monitor MLD snooping by using a variety of **show** commands.

In addition, you can clear the data for MLD groups and statistics by using **clear** commands. A **debug** command is also available. For command details, refer to the *Extreme SLX-OS Command Reference* for Extreme SLX 9140 and 9240 switches.

The following table lists the available **show** commands for MLD snooping.

TABLE 8 MLD snooping show commands

Command	Description
show ipv6 mld groups	Displays information about IPv6 MLDv1 groups.
show ipv6 mld snooping	Displays IPv6 MLD snooping details.
show ipv6 mld statistics	Displays IPv6 MLDv1 statistics.

The following table lists the available **clear** and **debug** commands.

TABLE 9 MLD snooping clear and debug commands

Command	Description
clear ipv6 mld groups	Clears IPv6 MLDv1 group cache entries.
clear ipv6 mld statistics	Clears IPv6 MLDv1 snooping statistics.
debug ipv6 mld	Displays information related to IPv6 MLD, with a variety of options.