Inspur

CN93240YC-FX2

NX-OS Multicast Routing

Configuration Guide

(Release 9.3.x)

Inspur-Cisco Networking Technology Co.,Ltd. provides customers with comprehensive technical support and services. For any assistance, please contact our local office or company headquarters.
Website: http://www.inspur.com/
Technical Support Tel: 400-691-1766
Technical Support Email:inspur_network@inspur.com
Technical Document Support Email:inspur_network@inspur.com
Address: 1036 Langchao Road, Lixia District, Jinan City, Shandong Province
Postal code: 250101

Preface

Objectives

This guide describes main functions of the CN93240YC-FX2. To have a quick grasp of the CN93240YC-FX2, please read this manual carefully.

Versions

The following table lists the product versions related to this document.

Product name	Version
CN93240YC-FX2	

Conventions

Symbol conventions

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

Symbol	Description
	Indicates a hazard with a medium or low level of risk which, if not avoided, could result in minor or moderate injury.
	Indicates a potentially hazardous situation that, if not avoided, could cause equipment damage, data loss, and performance degradation, or unexpected results.
	Provides additional information to emphasize or supplement important points of the main text.
	Indicates a tip that may help you solve a problem or save time.

General conventions

Convention	Description
Boldface	Names of files, directories, folders, and users are in **boldface**. For example, log in as user **root**.
Italic	Book titles are in *italics*.
`Lucida Console`	Terminal display is in `Lucida Console`.

Command conventions

Convention	Description
Boldface	The keywords of a command line are in **boldface**.
Italic	Command arguments are in *italics*.
[]	Items (keywords or arguments) in square brackets [ ] are optional.
{ x | y | ... }	Alternative items are grouped in braces and separated by vertical bars. One is selected.
[x | y | ...]	Optional alternative items are grouped in square brackets and separated by vertical bars. One or none is selected.
{ x | y | ... } *	Alternative items are grouped in braces and separated by vertical bars. A minimum of one or a maximum of all can be selected.
[x | y | ...] *	The parameter before the & sign can be repeated 1 to n times.

GUI conventions

Convention	Description
Boldface	Buttons, menus, parameters, tabs, windows, and dialog titles are in **boldface**. For example, click **OK**.
>	Multi-level menus are in boldface and separated by the ">" signs. For example, choose **File** > **Create** > **Folder**.

Keyboard operation

Format	Description
Key	Press the key. For example, press **Enter** and press **Tab**.

Format	Description
Key 1+Key 2	Press the keys concurrently. For example, pressing **Ctrl+C** means the two keys should be pressed concurrently.
Key 1, Key 2	Press the keys in turn. For example, pressing **Alt**, **A** means the two keys should be pressed in turn.

Mouse operation

Action	Description
Click	Select and release the primary mouse button without moving the pointer.
Double-click	Press the primary mouse button twice continuously and quickly without moving the pointer.
Drag	Press and hold the primary mouse button and move the pointer to a certain position.

Change history

Updates between document versions are cumulative. Therefore, the latest document version contains all updates made to previous versions.

Issue 01 (2020-02-24)

Initial commercial release

CONTENTS

Preface

This preface includes the following sections:

Audience

This publication is for network administrators who install, configure, and maintain CN switches.

Document Conventions

Command descriptions use the following conventions:

Convention	Description
bold	Bold text indicates the commands and keywords that you enter literally as shown.
Italic	Italic text indicates arguments for which you supply the values.
[x]	Square brackets enclose an optional element (keyword or argument).
[x | y]	Square brackets enclosing keywords or arguments that are separated by a vertical bar indicate an optional choice.
{x | y}	Braces enclosing keywords or arguments that are separated by a vertical bar indicate a required choice.
[x {y | z}]	Nested set of square brackets or braces indicate optional or required choices within optional or required elements. Braces and a vertical bar within square brackets indicate a required choice within an optional element.

Convention	Description
`variable`	Indicates a variable for which you supply values, in context where italics cannot be used.
string	A nonquoted set of characters. Do not use quotation marks around the string or the string includes the quotation marks.

Examples use the following conventions:

Convention	Description
`screen font`	Terminal sessions and information the switch displays are in screen font.
`boldface screen font`	Information that you must enter is in boldface screen font.
italic screen font	Arguments for which you supply values are in italic screen font.
< >	Nonprinting characters, such as passwords, are in angle brackets.
[ ]	Default responses to system prompts are in square brackets.
!, #	An exclamation point (!) or a pound sign (#) at the beginning of a line of code indicates a comment line.

Documentation Feedback

To provide technical feedback on this document, or to report an error or omission, please send your comments to inspur_network@inspur.com. We appreciate your feedback.

CHAPTER 1

New and Changed Information

This chapter provides release-specific information for each new and changed feature in the *CN93240YC-FX2 NX-OS Multicast Routing Configuration Guide, Release 9.3(x)*.

New and Changed Information

This table summarizes the new and changed features for the *CN93240YC-FX2 NX-OS Multicast Routing Configuration Guide, Release 9.3(x)* and where they are documented.

Table 1: New and Changed Features

Feature	Description	Changed in Release	Where Documented
IPv6 MLD Snooping	Added support for this feature.	9.3(3)	Configuring MLD Snooping
PIM6 support for SVI	Added support for this feature.	9.3(3)	Configuring PIM and PIM6
Multicast over GRE	Added support for this feature.	9.3(1)	Guidelines and Limitations for PIM and PIM6

New and Changed Information

CHAPTER 2

Platform Support for Multicast Routing Features

The following table describes platform support for features that are not supported across the entire suite of platforms. You should refer to each release's installation guide and release notes for details about the platforms supported in the initial product release.

Platform Support for Multicast Routing Features

The following table describes platform support for features that are not supported across the entire suite of Cisco Platforms. You should refer to each release's installation guide and release notes for details about the platforms supported in the initial product release.

IGMP

Return to Configuring IGMP, on page 19.

Feature	Supported Platform(s) or Line Cards	First Supported Release	Platform Exceptions
IGMP	CN93240YC-FX2 switches	7.0(3)I3(1)	None

PIM and PIM6

Return to Configuring PIM and PIM6, on page 55.

Feature	Supported Platform(s) or Line Cards	First Supported Release	Platform Exceptions
PIM and PIM6	CN93240YC-FX2 switches	7.0(3)I1(1)	None
PIM SSM and PIM ASM	CN93240YC-FX2 switches	7.0(3)I5(2)	None

IGMP Snooping

Return to Configuring IGMP Snooping.

Feature	Supported Platform(s) or Line Cards	First Supported Release	Platform Exceptions
IGMP Snooping	CN93240YC-FX2 switches	7.0(3)I4(2)	None

MSDP

Return to Configuring MSDP.

Feature	Supported Platform(s) or Line Cards	First Supported Release	Platform Exceptions
MSDP	CN93240YC-FX2 switches	-	None

NLB

Return to Configuring Microsoft Network Load Balancing (NLB).

Feature	Supported Platform(s) or Line Cards	First Supported Release	Platform Exceptions
NLB	CN93240YC-FX2 switches	9.2(1)	None

Platform Support for Multicast Routing Features

CHAPTER 3

Overview

This chapter describes the multicast features of NX-OS.

About Multicast

IP multicast is a method of forwarding the same set of IP packets to a number of hosts within a network. You can use multicast in IPv4 networks to provide efficient delivery of data to multiple destinations.

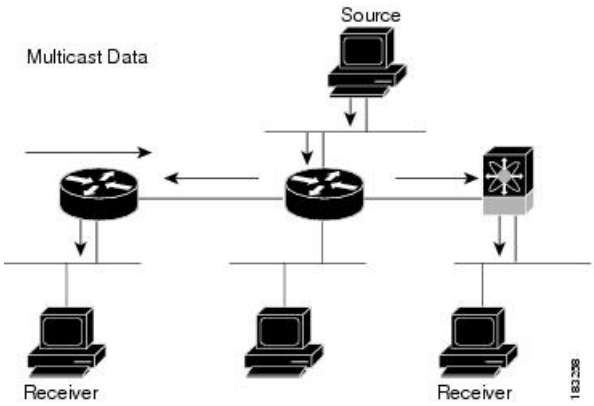
Multicast involves both a method of delivery and discovery of senders and receivers of multicast data, which is transmitted on IP multicast addresses called groups. A multicast address that includes a group and source IP address is often referred to as a channel. The Internet Assigned Number Authority (IANA) has assigned 224.0.0.0 through 239.255.255.255 as IPv4 multicast addresses.

Note For a complete list of RFCs related to multicast, see the *IETF RFCs for IP Multicast* chapter.

The routers in the network listen for receivers to advertise their interest in receiving multicast data from selected groups. The routers then replicate and forward the data from sources to the interested receivers. Multicast data for a group is transmitted only to those LAN segments with receivers that requested it.

This figure shows one source transmitting multicast data that is delivered to two receivers. In the figure, because the center host is on a LAN segment where no receiver requested multicast data, no data is delivered to that receiver.

Figure 1: Multicast Traffic from One Source to Two Receivers



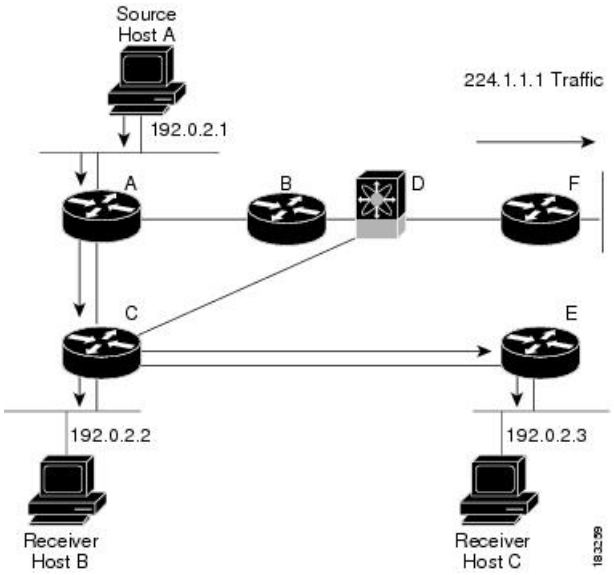
Multicast Distribution Trees

A multicast distribution tree represents the path that multicast data takes between the routers that connect sources and receivers. The multicast software builds different types of trees to support different multicast methods.

Source Trees

A source tree represents the shortest path that the multicast traffic takes through the network from the sources that transmit to a particular multicast group to receivers that requested traffic from that same group. Because of the shortest path characteristic of a source tree, this tree is often referred to as a shortest path tree (SPT). This figure shows a source tree for group 224.1.1.1 that begins at host A and connects to hosts B and C.

Figure 2: Source Tree

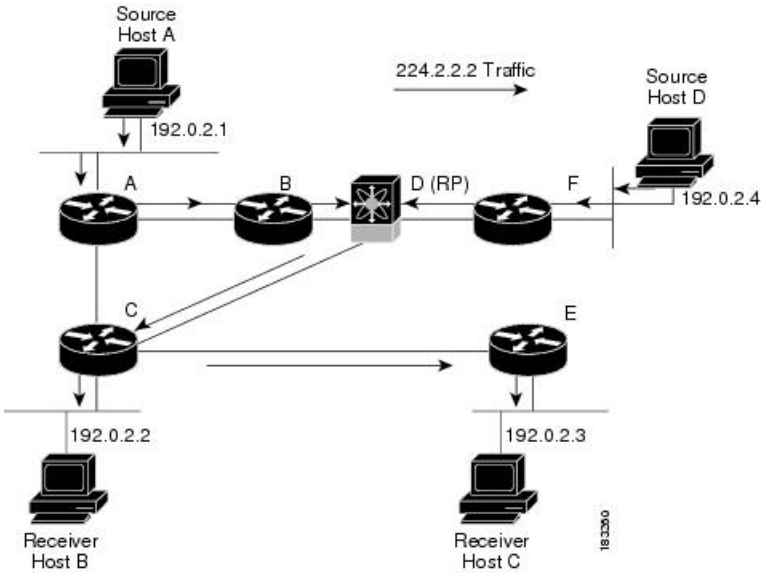


The notation (S, G) represents the multicast traffic from source S on group G. The SPT in this figure is written (192.0.2.1, 224.1.1.1). Multiple sources can be transmitting on the same group.

Shared Trees

A shared tree represents the shared distribution path that the multicast traffic takes through the network from a shared root or rendezvous point (RP) to each receiver. (The RP creates an SPT to each source.) A shared tree is also called an RP tree (RPT). This figure shows a shared tree for group 224.1.1.1 with the RP at router D. Source hosts A and D send their data to router D, the RP, which then forwards the traffic to receiver hosts B and C.

Figure 3: Shared Tree

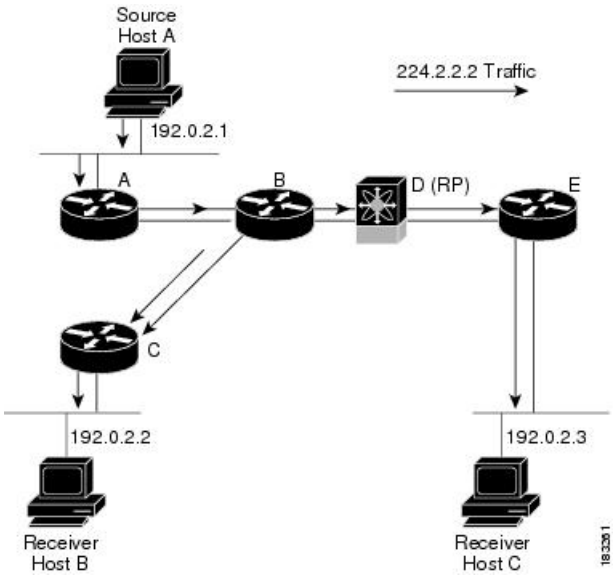


The notation (*, G) represents the multicast traffic from any source on group G. The shared tree in this figure is written (*, 224.2.2.2).

Bidirectional Shared Trees

A bidirectional shared tree represents the shared distribution path that the multicast traffic takes through the network from a shared root, or rendezvous point (RP), to each receiver. Multicast data is forwarded to receivers encountered on the way to the RP. The advantage of the bidirectional shared tree is shown in the figure below. Multicast traffic flows directly from host A to host B through routers B and C. In a shared tree, the data from source host A is first sent to the RP (router D) and then forwarded to router B for delivery to host B.

Figure 4: Bidirectional Shared Tree



The notation (*, G) represents the multicast traffic from any source on group G. The bidirectional tree in the figure is written as (*, 224.2.2.2).

Multicast Forwarding

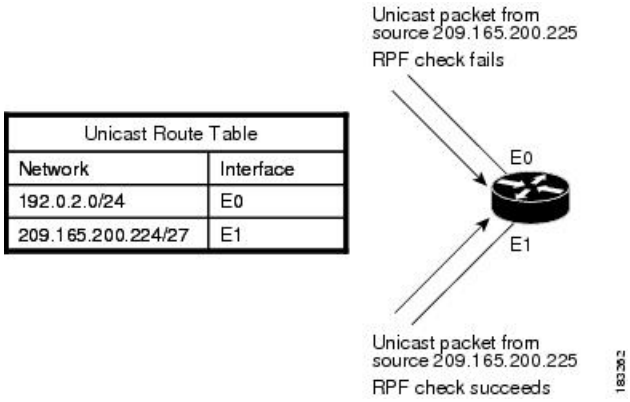
Because multicast traffic is destined for an arbitrary group of hosts, the router uses reverse path forwarding (RPF) to route data to active receivers for the group. When receivers join a group, a path is formed toward the source (SSM mode) or the RP (ASM or Bidir mode). The path from a source to a receiver flows in the reverse direction from the path that was created when the receiver joined the group.

For each incoming multicast packet, the router performs an RPF check. If the packet arrives on the interface leading to the source, the packet is forwarded out each interface in the outgoing interface (OIF) list for the group. Otherwise, the router drops the packet.

Note In Bidir mode, if a packet arrives on a non-RPF interface and the interface was elected as the designated forwarder (DF), then the packet is also forwarded in the upstream direction toward the RP.

This figure shows an example of RPF checks on packets coming in from different interfaces. The packet that arrives on E0 fails the RPF check because the unicast route table lists the source of the network on interface E1. The packet that arrives on E1 passes the RPF check because the unicast route table lists the source of that network on interface E1.

Figure 5: RPF Check Example



NX-OS PIM

NX-OS supports multicasting with Protocol Independent Multicast (PIM) sparse mode. PIM is IP routing protocol independent and can leverage whichever unicast routing protocols are used to populate the unicast routing table. In PIM sparse mode, multicast traffic is sent only to locations of the network that specifically request it. PIM dense mode is not supported by NX-OS.

Note In this publication, the term “PIM” is used for PIM sparse mode version 2.

To access multicast commands, you must enable the PIM feature. Multicast is enabled only after you enable PIM on an interface of each router in a domain. You can configure PIM for an IPv4 network. By default, IGMP is running on the system.

PIM, which is used between multicast-capable routers, advertises group membership across a routing domain by constructing multicast distribution trees. PIM builds shared distribution trees, on which packets from multiple sources are forwarded, as well as source distribution trees, on which packets from a single source are forwarded.

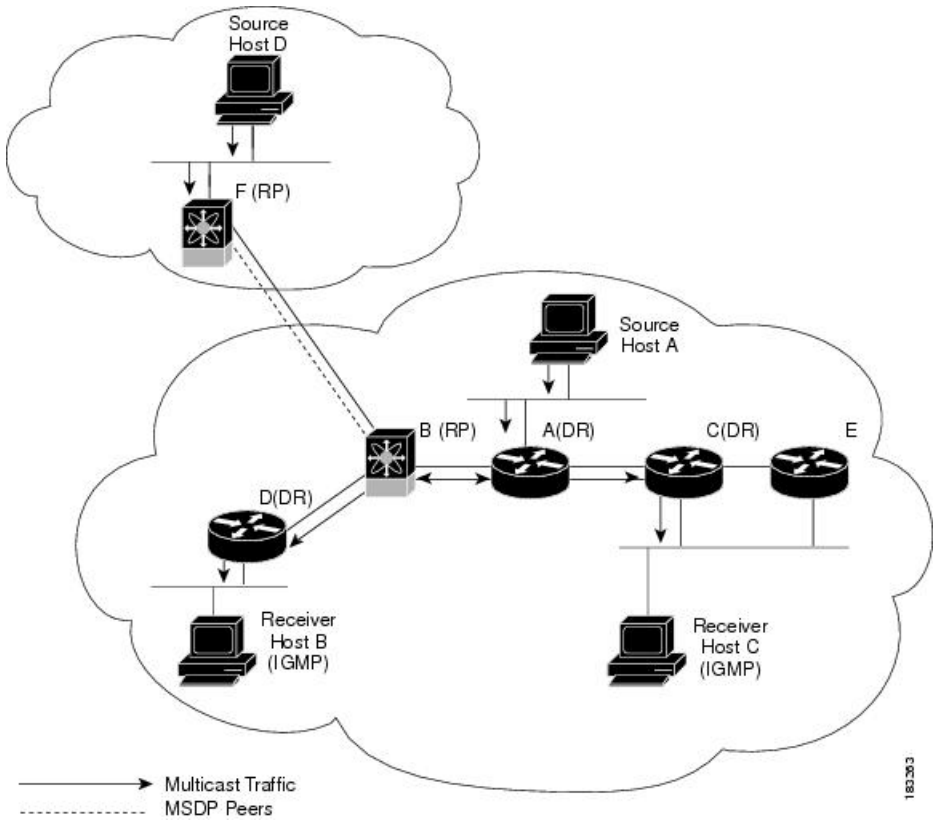
The distribution trees change automatically to reflect the topology changes due to link or router failures. PIM dynamically tracks both multicast-capable sources and receivers, although the source state is not created in Bidir mode.

The router uses the unicast routing table and RPF routes for multicast to create multicast routing information. In Bidir mode, additional multicast routing information is created.

Note In this publication, “PIM for IPv4” refers to the NX-OS implementation of PIM sparse mode.

This figure shows two PIM domains in an IPv4 network.

Figure 6: PIM Domains in an IPv4 Network



- The lines with arrows show the path of the multicast data through the network. The multicast data originates from the sources at hosts A and D.
- The dashed line connects routers B and F, which are Multicast Source Discovery Protocol (MSDP) peers. MSDP supports the discovery of multicast sources in other PIM domains.
- Hosts B and C receive multicast data by using Internet Group Management Protocol (IGMP) to advertise requests to join a multicast group.
- Routers A, C, and D are designated routers (DRs). When more than one router is connected to a LAN segment, such as C and E, the PIM software chooses one router to be the DR so that only one router is responsible for putting multicast data on the segment.

Router B is the rendezvous point (RP) for one PIM domain, and router F is the RP for the other PIM domain. The RP provides a common point for connecting sources and receivers within a PIM domain.

PIM supports these multicast modes for connecting sources and receivers:

- Any source multicast (ASM)
- Source-Specific Multicast (SSM)
- Bidirectional shared trees (Bidir)

NX-OS supports a combination of these modes for different ranges of multicast groups. You can also define RPF routes for multicast.

ASM

Any Source Multicast (ASM) is a PIM tree building mode that uses shared trees to discover new sources and receivers as well as source trees to form shortest paths from receivers to sources. The shared tree uses a network node as the root, called the rendezvous point (RP). The source tree is rooted at first-hop routers, directly attached to each source that is an active sender. The ASM mode requires an RP for a group range. An RP can be configured statically or learned dynamically by the Auto-RP or BSR group-to-RP discovery protocols. If an RP is learned and is not known to be a Bidir-RP, the group operates in ASM mode.

The ASM mode is the default mode when you configure RPs.

Bidir

Bidirectional shared trees (Bidir) is a PIM mode that, like the ASM mode, builds a shared tree between receivers and the RP but does not support switching over to a source tree when a new receiver is added to a group. In the Bidir mode, the router that is connected to a receiver is called the designated forwarder (DF) because multicast data can be forwarded directly from the designated router (DR) to the receiver without first going to the RP. The Bidir mode requires that you configure an RP.

The Bidir mode can reduce the amount of resources required on a router when there are many multicast sources and can continue to operate whether or not the RP is operational or connected.

SSM

Source-Specific Multicast (SSM) is a PIM mode that builds a source tree that originates at the designated router on the LAN segment that receives a request to join a multicast source. Source trees are built by sending PIM join messages in the direction of the source. The SSM mode does not require any RP configuration.

The SSM mode allows receivers to connect to sources outside the PIM domain.

RPF Routes for Multicast

You can configure static multicast RPF routes to override what the unicast routing table uses. This feature is used when the multicast topology is different than the unicast topology.

IGMP

By default, the Internet Group Management Protocol (IGMP) for PIM is running on the system.

IGMP is used by hosts that want to receive multicast data to request membership in multicast groups. Once the group membership is established, multicast data for the group is directed to the LAN segment of the requesting host.

You can configure IGMPv2 or IGMPv3 on an interface. You have to configure IGMPv3 with (S, G) to support SSM mode. By default, the software enables IGMPv2.

IGMP Snooping

IGMP snooping is a feature that limits multicast traffic on VLANs to the subset of ports that have known receivers. By examining (snooping) IGMP membership report messages from interested hosts, multicast traffic is sent only to VLAN ports that interested hosts reside on. By default, IGMP snooping is running on the system.

Interdomain Multicast

NX-OS provides several methods that allow multicast traffic to flow between PIM domains.

SSM

The PIM software uses SSM to construct a shortest path tree from the designated router for the receiver to a known source IP address, which may be in another PIM domain. The ASM and Bidir modes mode cannot access sources from another PIM domain without the use of another protocol.

Once you enable PIM in your networks, you can use SSM to reach any multicast source that has an IP address known to the designated router for the receiver.

MSDP

Multicast Source Discovery Protocol (MSDP) is a multicast routing protocol that is used with PIM to support the discovery of multicast sources in different PIM domains.

Note NX-OS supports the PIM Anycast-RP, which does not require MSDP configuration.

MBGP

Multiprotocol BGP (MBGP) defines extensions to BGP4 that enable routers to carry multicast routing information. PIM can use this multicast information to reach sources in external BGP autonomous systems.

MRIB

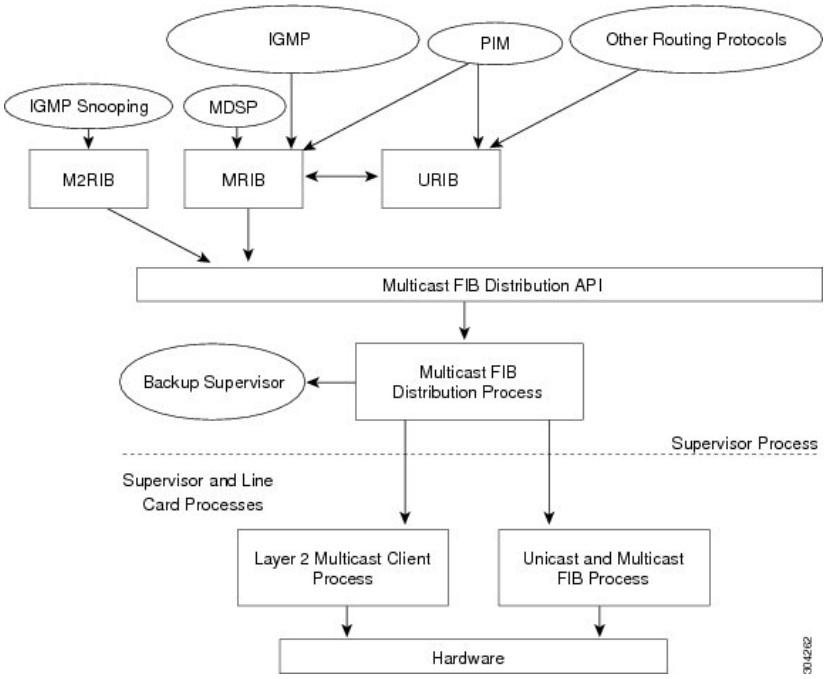
The NX-OS IPv4 Multicast Routing Information Base (MRIB) is a repository for route information that is generated by multicast protocols such as PIM and IGMP. The MRIB does not affect the route information itself. The MRIB maintains independent route information for each virtual routing and forwarding (VRF) instance.

The major components of the NX-OS multicast software architecture are as follows:

- The Multicast FIB (MFIB) Distribution (MFDM) API defines an interface between the multicast Layer 2 and Layer 3 control plane modules, including the MRIB, and the platform forwarding plane. The control plane modules send the Layer 3 route update using the MFDM API.
- The multicast FIB distribution process distributes the multicast update messages to all the relevant modules and the standby supervisor. It runs only on the supervisor.
- The Layer 2 multicast client process sets up the Layer 2 multicast hardware forwarding path. It runs on both the supervisor and the modules.
- The unicast and multicast FIB process manages the Layer 3 hardware forwarding path. It runs on both the supervisor and the modules.

  The following figure shows the NX-OS multicast software architecture.=

Figure 7: NX-OS Multicast Software Architecture



Virtual Port Channels and Multicast

A virtual port channel (vPC) allows a single device to use a port channel across two upstream switches. When you configure a vPC, the following multicast features might be affected:

- PIM—NX-OS software for the CN93240YC-FX2 switches does not support PIM Bidir on a vPC.
- IGMP snooping—You should configure the vPC peers identically.=

Licensing Requirements for Multicast

The multicast features that require a license are as follows:

- PIM
- MSDP

The multicast features that require no license are as follows:

- IGMP
- IGMP snooping

For a complete explanation of the NX-OS licensing scheme, see *NX-OS Licensing Guide*.

Guidelines and Limitations for Multicast

- Layer 3 Ethernet port-channel subinterfaces are not supported with multicast routing.
- Layer 2 IPv6 multicast packets will be flooded on the incoming VLAN.
- Traffic storm control is not supported for unknown multicast traffic.=

High-Availability Requirements for Multicast

After a multicast routing protocol is restarted, its state is recovered from the MRIB process. When a supervisor switchover occurs, the MRIB recovers its state from the hardware, and the multicast protocols recover their state from periodic message activity. For more information about high availability, see the *CN93240YC-FX2 NX-OS High Availability and Redundancy Guide*.

Virtual Device Contexts

NX-OS can segment operating system and hardware resources into virtual device contexts (VDCs) that emulate virtual devices. The CN93240YC-FX2 switches currently do not support multiple VDCs. All switch resources are managed in the default VDC.

Troubleshooting Inconsistency Between SW and HW Multicast Routes

Symptom

This section provides symptoms, possible causes, and recommended actions for when *, G, or S,G entries that are seen in the MRIB with active flow, but are not programmed in MFIB.

Possible Cause

The issue can be seen when numerous active flows are received beyond the hardware capacity. This causes some of the entries not to be programmed in hardware while there is no free hardware index.

If the number of active flows are significantly reduced to free up the hardware resource, inconsistency may be seen between MRIB and MFIB for flows that were previously affected when the hardware table was full until the entry, times out, repopulates, and triggers programming.

There is currently no mechanism to walk the MRIB table and reprogram missing entries in HW after hardware resource is freed.

Corrective Action

To ensure reprogramming of the entries, use the **clear ip mroute *** command.

CHAPTER 4

Configuring IGMP

This chapter describes how to configure the Internet Group Management Protocol (IGMP) on NX-OS devices for IPv4 networks.

About IGMP

IGMP is an IPv4 protocol that a host uses to request multicast data for a particular group. Using the information obtained through IGMP, the software maintains a list of multicast group or channel memberships on a per-interface basis. The systems that receive these IGMP packets send multicast data that they receive for requested groups or channels out the network segment of the known receivers.

By default, the IGMP process is running. You cannot enable IGMP manually on an interface. IGMP is automatically enabled when you perform one of the following configuration tasks on an interface:

- Enable PIM
- Statically bind a local multicast group
- Enable link-local group reports

IGMP Versions

The device supports IGMPv2 and IGMPv3, and IGMPv1 report reception.

By default, the software enables IGMPv2 when it starts the IGMP process. You can enable IGMPv3 on interfaces where you want its capabilities.

IGMPv3 includes the following key changes from IGMPv2:

- Support for Source-Specific Multicast (SSM), which builds shortest path trees from each receiver to the source, through the following features:
  - Host messages that can specify both the group and the source.
  - The multicast state that is maintained for groups and sources, not just for groups as in IGMPv2.
- Hosts no longer perform report suppression, which means that hosts always send IGMP membership reports when an IGMP query message is received.

Note The CN93240YC-FX2 switches do not support SSM until NX-OS Release 7.0(3)I2(1).

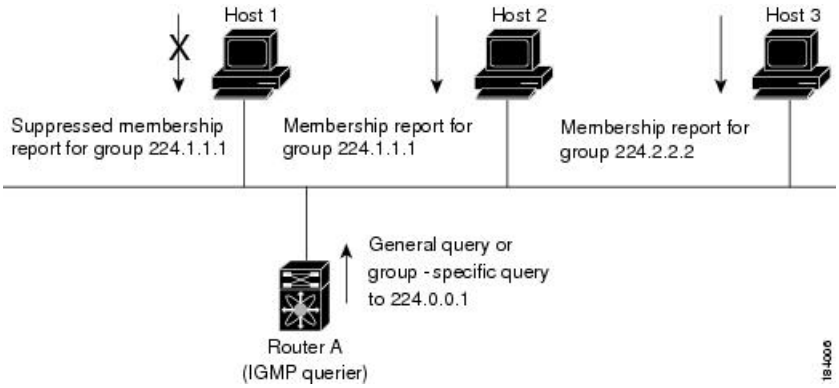
For detailed information about IGMPv2, see RFC 2236.

For detailed information about IGMPv3, see RFC 3376.

IGMP Basics

This figure shows the basic IGMP process of a router that discovers multicast hosts. Hosts 1, 2, and 3 send unsolicited IGMP membership report messages to initiate receiving multicast data for a group or channel.

Figure 8: IGMPv1 and IGMPv2 Query-Response Process



In the figure below, router A, which is the IGMP designated querier on the subnet, sends query messages to the all-hosts multicast group at 224.0.0.1 periodically to discover whether any hosts want to receive multicast data. You can configure the group membership timeout value that the router uses to determine that no members of a group or source exist on the subnet.

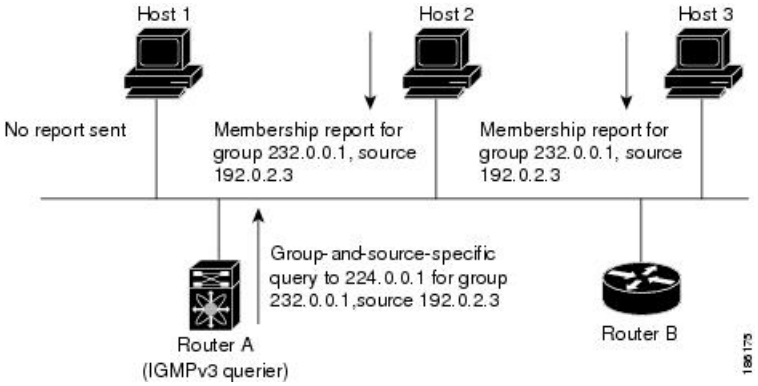
The software elects a router as the IGMP querier on a subnet if it has the lowest IP address. As long as a router continues to receive query messages from a router with a lower IP address, it resets a timer that is based on its querier timeout value. If the querier timer of a router expires, it becomes the designated querier. If that router later receives a host query message from a router with a lower IP address, it drops its role as the designated querier and sets its querier timer again.

In this figure, host 1's membership report is suppressed, and host 2 sends its membership report for group 224.1.1.1 first. Host 1 receives the report from host 2. Because only one membership report per group needs to be sent to the router, other hosts suppress their reports to reduce network traffic. Each host waits for a random time interval to avoid sending reports at the same time. You can configure the query maximum response time parameter to control the interval in which hosts randomize their responses.

Note IGMPv1 and IGMPv2 membership report suppression occurs only on hosts that are connected to the same port.

In this figure, router A sends the IGMPv3 group-and-source-specific query to the LAN. Hosts 2 and 3 respond to the query with membership reports that indicate that they want to receive data from the advertised group and source. This IGMPv3 feature supports SSM.

Figure 9: IGMPv3 Group-and-Source-Specific Query



Note IGMPv3 hosts do not perform IGMP membership report suppression.

Messages sent by the designated querier have a time-to-live (TTL) value of 1, which means that the messages are not forwarded by the directly connected routers on the subnet. You can configure the frequency and number of query messages sent specifically for IGMP startup, and you can configure a short query interval at startup so that the group state is established as quickly as possible. Although usually unnecessary, you can tune the query interval used after startup to a value that balances the responsiveness to host group membership messages and the traffic created on the network.

Caution Changing the query interval can severely impact multicast forwarding.

When a multicast host leaves a group, a host that runs IGMPv2 or later sends an IGMP leave message. To check if this host is the last host to leave the group, the software sends an IGMP query message and starts a timer that you can configure called the last member query response interval. If no reports are received before the timer expires, the software removes the group state. The router continues to send multicast traffic for a group until its state is removed.

You can configure a robustness value to compensate for packet loss on a congested network. The robustness value is used by the IGMP software to determine the number of times to send messages.

Link local addresses in the range 224.0.0.0/24 are reserved by the Internet Assigned Numbers Authority (IANA). Network protocols on a local network segment use these addresses; routers do not forward these addresses because they have a TTL of 1. By default, the IGMP process sends membership reports only for nonlink local addresses, but you can configure the software to send reports for link local addresses.

Licensing Requirements for IGMP

Product	License Requirement
NX-OS	IGMP requires no license. Any feature not included in a license package is bundled with the nx-os image and is provided at no extra charge to you. For a complete explanation of the NX-OS licensing scheme, see the *NX-OS Licensing Guide*.

Prerequisites for IGMP

IGMP has the following prerequisites:

- You are logged onto the device.
- For global configuration commands, you are in the correct virtual routing and forwarding (VRF) mode. The default configuration mode shown in the examples in this chapter applies to the default VRF.

Guidelines and Limitations for IGMP

IGMP has the following guidelines and limitations:

- Excluding or blocking a list of sources according to IGMPv3 (RFC 3376) is not supported.=

Default Settings for IGMP

This table lists the default settings for IGMP parameters.

Table 2: Default IGMP Parameters

Parameters	Default
IGMP version	2
Startup query interval	30 seconds
Startup query count	2
Robustness value	2

Parameters	Default
Querier timeout	255 seconds
Query timeout	255 seconds
Query max response time	10 seconds
Query interval	125 seconds
Last member query response interval	1 second
Last member query count	2
Group membership timeout	260 seconds
Report link local multicast groups	Disabled
Enforce router alert	Disabled
Immediate leave	Disabled

Configuring IGMP Parameters

You can configure the IGMP global and interface parameters to affect the operation of the IGMP process.

Note If you are familiar with the IOS CLI, be aware that the NX-OS commands for this feature might differ from the IOS commands that you would use.

Configuring IGMP Interface Parameters

You can configure the optional IGMP interface parameters described in the table below.

Table 3: IGMP Interface Parameters

Parameter	Description
IGMP version	IGMP version that is enabled on the interface. The IGMP version can be 2 or 3. The default is 2.

Parameter	Description
Static multicast groups	Multicast groups that are statically bound to the interface. You can configure the groups to join the interface with the (*, G) state or specify a source IP to join with the (S, G) state. You can specify a route-map policy name that lists the group prefixes, group ranges, and source prefixes to use with the **match ip multicast** command. **Note** Although you can configure the (S, G) state, the source tree is built only if you enable IGMPv3. You can configure a multicast group on all the multicast-capable routers on the network so that pinging the group causes all the routers to respond.
Static multicast groups on OIF	Multicast groups that are statically bound to the output interface. You can configure the groups to join the output interface with the (*, G) state or specify a source IP to join with the (S, G) state. You can specify a route-map policy name that lists the group prefixes, group ranges, and source prefixes to use with the **match ip multicast** command. **Note** Although you can configure the (S, G) state, the source tree is built only if you enable IGMPv3.
Startup query interval	Startup query interval. By default, this interval is shorter than the query interval so that the software can establish the group state as quickly as possible. Values range from 1 to 18,000 seconds. The default is 31 seconds.
Startup query count	Number of queries sent at startup that are separated by the startup query interval. Values range from 1 to 10. The default is 2.
Robustness value	Robustness variable that you can tune to reflect expected packet loss on a congested network. You can increase the robustness variable to increase the number of times that packets are resent. Values range from 1 to 7. The default is 2.
Querier timeout	Number of seconds that the software waits after the previous querier has stopped querying and before it takes over as the querier. Values range from 1 to 65,535 seconds. The default is 255 seconds.

Parameter	Description
Query max response time	Maximum response time advertised in IGMP queries. You can tune the IGMP messages on the network by setting a larger value so that host responses are spread out over a longer time. This value must be less than the query interval. Values range from 1 to 25 seconds. The default is 10 seconds.
Query interval	Frequency at which the software sends IGMP host query messages. You can tune the number of IGMP messages on the network by setting a larger value so that the software sends IGMP queries less often. Values range from 1 to 18,000 seconds. The default is 125 seconds.
Last member query response interval	Interval in which the software sends a response to an IGMP query after receiving a host leave message from the last known active host on the subnet. If no reports are received in the interval, the group state is deleted. You can use this value to tune how quickly the software stops transmitting on the subnet. The software can detect the loss of the last member of a group or source more quickly when the values are smaller. Values range from 1 to 25 seconds. The default is 1 second.
Last member query count	Number of times that the software sends an IGMP query, separated by the last member query response interval, in response to a host leave message from the last known active host on the subnet. Values range from 1 to 5. The default is 2. Setting this value to 1 means that a missed packet in either direction causes the software to remove the multicast state from the queried group or channel. The software may wait until the next query interval before the group is added again.
Group membership timeout	Group membership interval that must pass before the router decides that no members of a group or source exist on the network. Values range from 3 to 65,535 seconds. The default is 260 seconds.
Report link local multicast groups	Option that enables sending reports for groups in 224.0.0.0/24. Link local addresses are used only by protocols on the local network. Reports are always sent for nonlink local groups. The default is disabled.
Report policy	Access policy for IGMP reports that is based on a route-map policy. 1

Parameter	Description
Access groups	Option that configures a route-map policy to control the multicast groups that hosts on the subnet serviced by an interface can join. **Note** Only the **match ip multicast group** command is supported in this route map policy. The **match ip address** command for matching an ACL is not supported.
Immediate leave	Option that minimizes the leave latency of IGMPv2 group memberships on a given IGMP interface because the device does not send group-specific queries. When immediate leave is enabled, the device removes the group entry from the multicast routing table immediately upon receiving a leave message for the group. The default is disabled. **Note** Use this command only when there is one receiver behind the interface for a given group.

[1] To configure route-map policies, see the *CN93240YC-FX2 NX-OS Unicast Routing Configuration Guide*.

Procedure

	Command or Action	Purpose

Step 1	**configure terminal**
	Example:
	`switch# configure terminal`
	`switch(config)#`
		Enters global configuration mode.

Step 2	**interface** *interface*
	Example:
	`switch(config)# interface ethernet 2/1`
	`switch(config-if)#`
		Enters interface configuration mode.

Step 3	**Option**: **ip igmp version** *value* **Example:** `switch(config-if)# ip igmp version 3`	**Description**: Sets the IGMP version to the value specified. Values can be 2 or 3. The default is 2. The **no** form of the command sets the version to 2.

Command or Action	Purpose
Option	**Description**
ip igmp join-group {**group** [**source** *source*] | **route-map** *policy-name*} **Example:** `switch(config-if)# ip igmp join-group 230.0.0.0`	Configures an interface on the device to join the specified group or channel. The device accepts the multicast packets for CPU consumption only. **Caution** The device CPU must be able to handle the traffic generated by using this command. Because of CPU load constraints, using this command, especially in any form of scale, is not recommended. Consider using the **ip igmp static-oif** command instead.
ip igmp static-oif {*group* [**source** *source*] | **route-map** *policy-name*} **Example:** `switch(config-if)# ip igmp static-oif 230.0.0.0`	Statically binds a multicast group to the outgoing interface, which is handled by the device hardware. If you specify only the group address, the (*, G) state is created. If you specify the source address, the (S, G) state is created. You can specify a route-map policy name that lists the group prefixes, group ranges, and source prefixes to use with the **match ip multicast** command.

Command or Action	Purpose
Option	**Description**
	Note A source tree is built for the (S, G) state only if you enable IGMPv3.
ip igmp startup-query-interval *seconds* **Example:** `switch(config-if)# ip igmp startup-query-interval 25`	Sets the query interval used when the software starts up. Values can range from 1 to 18,000 seconds. The default is 31 seconds.
ip igmp startup-query-count *count* **Example:** `switch(config-if)# ip igmp startup-query-count 3`	Sets the query count used when the software starts up. Values can range from 1 to 10. The default is 2.
ip igmp robustness-variable *value* **Example:** `switch(config-if)# ip igmp robustness-variable 3`	Sets the robustness variable. Values can range from 1 to 7. The default is 2.
ip igmp querier-timeout *seconds* **Example:** `switch(config-if)# ip igmp querier-timeout 300`	Sets the querier timeout that the software uses when deciding to take over as the querier. Values can range from 1 to 65,535 seconds. The default is 255 seconds.
ip igmp query-timeout *seconds* **Example:** `switch(config-if)# ip igmp query-timeout 300`	Sets the query timeout that the software uses when deciding to take over as the querier. Values can range from 1 to 65,535 seconds. The default is 255 seconds. **Note** This command has the same functionality as the **ip igmp querier-timeout** command.

Command or Action	Purpose
Option	**Description**
ip igmp query-max-response-time *seconds* **Example:** `switch(config-if)# ip igmp query-max-response-time 15`	Sets the response time advertised in IGMP queries. Values can range from 1 to 25 seconds. The default is 10 seconds.
ip igmp query-interval *interval* **Example:** `switch(config-if)# ip igmp query-interval 100`	Sets the frequency at which the software sends IGMP host query messages. Values can range from 1 to 18,000 seconds. The default is 125 seconds.
ip igmp last-member-query-response-time *seconds* **Example:** `switch(config-if)# ip igmp last-member-query-response-time 3`	Sets the query interval waited after sending membership reports before the software deletes the group state. Values can range from 1 to 25 seconds. The default is 1 second.
ip igmp last-member-query-count *count* **Example:** `switch(config-if)# ip igmp last-member-query-count 3`	Sets the number of times that the software sends an IGMP query in response to a host leave message. Values can range from 1 to 5. The default is 2.
ip igmp group-timeout *seconds* **Example:** `switch(config-if)# ip igmp group-timeout 300`	Sets the group membership timeout for IGMPv2. Values can range from 3 to 65,535 seconds. The default is 260 seconds.
ip igmp report-link-local-groups **Example:** `switch(config-if)# ip igmp report-link-local-groups`	Enables sending reports for groups in 224.0.0.0/24. Reports are always sent for nonlink local groups. By default, reports are not sent for link local groups.
ip igmp report-policy *policy* **Example:**	Configures an access policy for IGMP reports

Command or Action	Purpose
Option	**Description**
`switch(config-if)# ip igmp report-policy my_report_policy`	that is based on a route-map policy.
ip igmp access-group *policy* **Example:** `switch(config-if)# ip igmp access-group my_access_policy`	Configures a route-map policy to control the multicast groups that hosts on the subnet serviced by an interface can join. **Note** Only the **match ip multicast group** command is supported in this route map policy. The **match ip address** command for matching an ACL is not supported.
ip igmp immediate-leave **Example:** `switch(config-if)# ip igmp immediate-leave`	Enables the device to remove the group entry from the multicast routing table immediately upon receiving a leave message for the group. Use this command to minimize the leave latency of IGMPv2 group memberships on a given IGMP interface because the device does not send group-specific queries. The default is disabled. **Note** Use this command only when there is one receiver behind the interface for a given group.

	Command or Action	Purpose
Step 4	(Optional) **show ip igmp interface** [*interface*] [**vrf** *vrf-name* | **all**] [**brief**] **Example:** `switch(config)# show ip igmp interface`	Displays IGMP information about the interface.
Step 5	(Optional) **copy running-config startup-config** **Example:** `switch(config)# copy running-config startup-config`	Copies the running configuration to the startup configuration.

Configuring an IGMP SSM Translation

You can configure an SSM translation to provide SSM support when the router receives IGMPv1 or IGMPv2 membership reports. Only IGMPv3 provides the capability to specify group and source addresses in membership reports. By default, the group prefix range is 232.0.0.0/8.

The IGMP SSM translation feature enables an SSM-based multicast core network to be deployed when the multicast host does not support IGMPv3 or is forced to send group joins instead of (S,G) reports to interoperate with Layer 2 switches. The IGMP SSM translation feature provides the functionality to configure multiple sources for the same SSM group. Protocol Independent Multicast (PIM) must be configured on the device before configuring the SSM translation.

This table lists the example SSM translations.

Table 4: Example SSM Translations

Group Prefix	Source Address
232.0.0.0/8	10.1.1.1
232.0.0.0/8	10.2.2.2
232.1.0.0/16	10.3.3.3
232.1.1.0/24	10.4.4.4

This table shows the resulting MRIB routes that the IGMP process creates when it applies an SSM translation to the IGMP membership report. If more than one translation applies, the router creates the (S, G) state for each translation.

Table 5: Example Result of Applying SSM Translations

IGMPv2 Membership Report	Resulting MRIB Route
232.1.1.1	(10.4.4.4, 232.1.1.1)
232.2.2.2	(10.1.1.1, 232.2.2.2) (10.2.2.2, 232.2.2.2)

Procedure

	Command or Action	Purpose

Step 1	**configure terminal**
	Example:
	`switch# configure terminal`
	`switch(config)#`
		Enters global configuration mode.

Step 2	**ip igmp ssm-translate** *group-prefix source-addr* **Example:** `switch(config)# ip igmp ssm-translate 232.0.0.0/8 10.1.1.1`	Configures the translation of IGMPv1 or IGMPv2 membership reports by the IGMP process to create the (S,G) state as if the router had received an IGMPv3 membership report.
Step 3	(Optional) **show running-configuration igmp** **Example:** `switch(config)# show running-configuration igmp`	Shows the running-configuration information, including **ssm-translate** command lines.
Step 4	(Optional) **copy running-config startup-config** **Example:** `switch(config)# copy running-config startup-config`	Copies the running configuration to the startup configuration.

Configuring the Enforce Router Alert Option Check

You can configure the enforce router alert option check for IGMPv2 and IGMPv3 packets.

Procedure

	Command or Action	Purpose

Step 1	**configure terminal**
	Example:
	`switch# configure terminal`
	`switch(config)#`
		Enters global configuration mode.

Step 2	[**no**] **ip igmp enforce-router-alert** **Example:** `switch(config)# ip igmp enforce-router-alert`	Enables or disables the enforce router alert option check for IGMPv2 and IGMPv3 packets. By default, the enforce router alert option check is enabled.
Step 3	(Optional) **show running-configuration igmp** **Example:** `switch(config)# show running-configuration igmp`	Shows the running-configuration information.

	Command or Action	Purpose
Step 4	(Optional) **copy running-config startup-config** **Example:** `switch(config)# copy running-config startup-config`	Copies the running configuration to the startup configuration.

Restarting the IGMP Process

You can restart the IGMP process and optionally flush all routes.

Procedure

	Command or Action	Purpose
Step 1	**restart igmp** **Example:** `switch# restart igmp`	Restarts the IGMP process.

Step 2	**configure terminal**
	Example:
	`switch# configure terminal`
	`switch(config)#`
		Enters global configuration mode.

Step 3	**ip igmp flush-routes** **Example:** `switch(config)# ip igmp flush-routes`	Removes routes when the IGMP process is restarted. By default, routes are not flushed.
Step 4	(Optional) **show running-configuration igmp** **Example:** `switch(config)# show running-configuration igmp`	Shows the running-configuration information.
Step 5	(Optional) **copy running-config startup-config** **Example:** `switch(config)# copy running-config startup-config`	Copies the running configuration to the startup configuration.

Verifying the IGMP Configuration

To display the IGMP configuration information, perform one of the following tasks:

Command	Description
show ip igmp interface [*interface*] [**vrf** *vrf-name* | **all**] [**brief**]	Displays IGMP information about all interfaces or a selected interface, the default VRF, a selected VRF, or all VRFs. If IGMP is in vPC mode, use this command to display vPC statistics.
show ip igmp groups [{**source** [*group*]}] | {**group** [*source*]}] [**interface**] [**summary**] [**vrf** *vrf-name* | **all**]	Displays the IGMP attached group membership for a group or interface, the default VRF, a selected VRF, or all VRFs.
show ip igmp route [{**source** [*group*]}] | {**group** [*source*]}] [**interface**] [**summary**] [**vrf** *vrf-name* | **all**]	Displays the IGMP attached group membership for a group or interface, the default VRF, a selected VRF, or all VRFs.
show ip igmp local-groups	Displays the IGMP local group membership.
show running-configuration igmp	Displays the IGMP running-configuration information.
show startup-configuration igmp	Displays the IGMP startup-configuration information.

Configuration Examples for IGMP

The following example shows how to configure the IGMP parameters:

```
configure terminal
  ip igmp ssm-translate 232.0.0.0/8 10.1.1.1
  interface ethernet 2/1
    ip igmp version 3
    ip igmp join-group 230.0.0.0
    ip igmp startup-query-interval 25
    ip igmp startup-query-count 3
    ip igmp robustness-variable 3
    ip igmp querier-timeout 300
    ip igmp query-timeout 300
    ip igmp query-max-response-time 15
    ip igmp query-interval 100
    ip igmp last-member-query-response-time 3
    ip igmp last-member-query-count 3
    ip igmp group-timeout 300
    ip igmp report-link-local-groups
    ip igmp report-policy my_report_policy
    ip igmp access-group my_access_policy
```

CHAPTER 5

Configuring MLD

This chapter describes how to configure Multicast Listener Discovery (MLD) on NX-OS devices for IPv6 networks.

About MLD

MLD is an IPv6 protocol that a host uses to request multicast data for a particular group. Using the information obtained through MLD, the software maintains a list of multicast group or channel memberships on a per-interface basis. The devices that receive MLD packets send the multicast data that they receive for requested groups or channels out the network segment of the known receivers.

MLDv1 is derived from IGMPv2, and MLDv2 is derived from IGMPv3. IGMP uses IP Protocol 2 message types while MLD uses IP Protocol 58 message types, which is a subset of the ICMPv6 messages.

The MLD process is started automatically on the device. You cannot enable MLD manually on an interface. MLD is enabled automatically when you perform one of the following configuration tasks on an interface:

- Enable PIM6
- Statically bind a local multicast group
- Enable link-local group reports

MLD Versions

The device supports MLDv1 and MLDv2. MLDv2 supports MLDv1 listener reports.

By default, the software enables MLDv2 when it starts the MLD process. You can enable MLDv1 on interfaces where you want only its capabilities.

MLDv2 includes the following key changes from MLDv1:

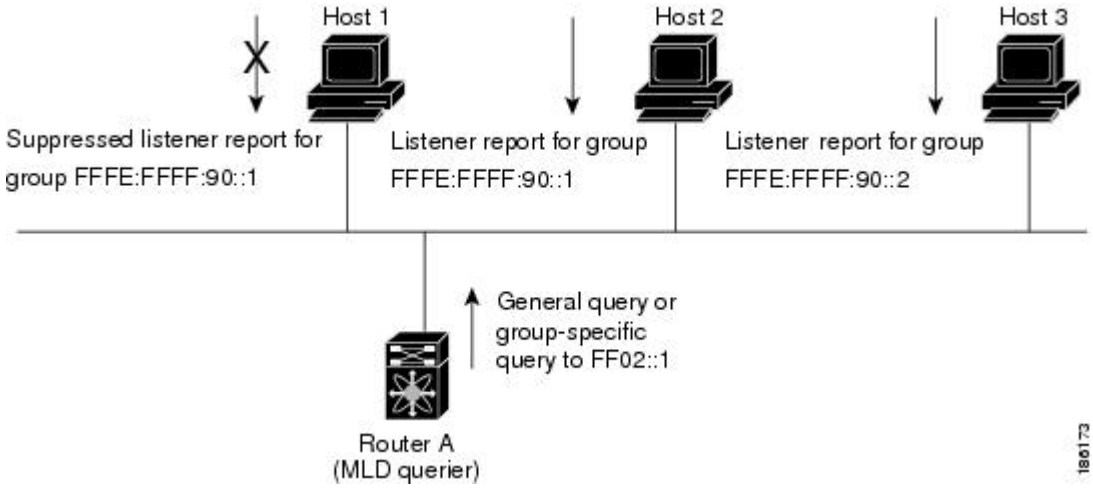
- Support for Source-Specific Multicast (SSM), which builds shortest path trees from each receiver to the source, through the following features:
  - Host messages that can specify both the group and the source.
  - The multicast state that is maintained for groups and sources, not just for groups as in MLDv1.
- Hosts no longer perform report suppression, which means that hosts always send MLD listener reports when an MLD query message is received.

For detailed information about MLDv1, see RFC 2710. For detailed information about MLDv2, see RFC 3810.

MLD Basics

The basic MLD process of a router that discovers multicast hosts is shown in the figure below.

Figure 10: MLD Query-Response Process



Hosts 1, 2, and 3 send unsolicited MLD listener report messages to initiate receiving multicast data for a group or channel. Router A, which is the MLD designated querier on the subnet, sends a general query message to the link-scope all-nodes multicast address FF02::1 periodically to discover which multicast groups hosts want to receive. The group-specific query is used to discover whether a specific group is requested by any hosts. You can configure the group membership timeout value that the router uses to determine if any members of a group or source exist on the subnet.

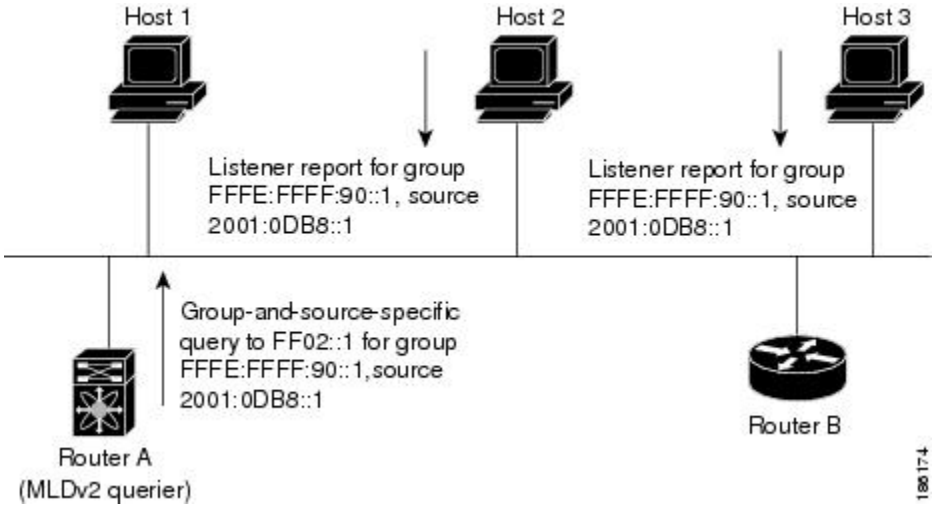
Host 1's listener report is suppressed, and host 2 sends its listener report for group FFFE:FFFF:90::1 first. Host 1 receives the report from host 2. Because only one listener report per group needs to be sent to the router, other hosts suppress their reports to reduce network traffic. Each host waits for a random time interval to avoid sending reports at the same time. You can configure the query maximum response time parameter to control the interval at which hosts randomize their responses.

Note MLDv1 membership report suppression occurs only on hosts that are connected to the same port.

Router A sends the MLDv2 group-and-source-specific query to the LAN. Hosts 2 and 3 respond to the query with listener reports to indicate that they want to receive data from the advertised group and source. This MLDv2 feature supports SSM.

Note In MLDv2, all hosts respond to queries.

Figure 11: MLDv2 Group-and-Source-Specific Query



The software elects a router as the MLD querier on a subnet if it has the lowest IP address. As long as a router continues to receive query messages from a router with a lower IP address, it remains a nonquerier and resets a timer that is based on its querier timeout value. If the querier timer of a router expires, it becomes the designated querier. If that router later receives a host query message from a router with a lower IP address, it drops its role as the designated querier and sets its querier timer again.

Messages sent by the designated querier have a time-to-live (TTL) value of 1, which means that the messages are not forwarded by the directly connected routers on the subnet, and you can configure the frequency and number of query messages sent specifically for MLD startup. You can configure a short query interval at startup so that the group state is established as quickly as possible. Although usually unnecessary, you can tune the query interval used after startup to a value that balances responsiveness to host group membership and the traffic created on the network.

Caution If you change the query interval, you can severely impact multicast forwarding in your network.

When a multicast host leaves a group, it should send a done message for MLDv1 or a listener report that excludes the group to the link-scope all-routers multicast address FF02::2. To check if this host is the last host to leave the group, the software sends an MLD query message and starts a timer that you can configure called the last member query response interval. If no reports are received before the timer expires, the software removes the group state. The router continues to send multicast traffic for a group until its state is removed.

You can configure a robustness value to compensate for the packet loss on a congested network. The robustness value is used by the MLD software to determine the number of times to send messages.

Link local addresses in the range FF02::0/16 have link scope, as defined by the Internet Assigned Numbers Authority (IANA). Network protocols on a local network segment use these addresses; routers do not forward these addresses because they have a TTL of 1. By default, the MLD process sends listener reports only for nonlink local addresses, but you can configure the software to send reports for link local addresses.

MLD Snooping

Multicast Listener Discovery (MLD) snooping enables the efficient distribution of IPv6 multicast traffic between hosts and routers. It is a Layer 2 feature that restricts IPv6 multicast traffic within a bridge-domain to a subset of ports that have transmitted or received MLD queries or reports. In this way, MLD snooping provides the benefit of conserving the bandwidth on those segments of the network where no node has expressed interest in receiving the multicast traffic. This reduces the bandwidth usage instead of flooding the bridge-domain, and also helps hosts and routers save unwanted packet processing.

The MLD snooping functionality is similar to Internet Group Management Protocol (IGMP) snooping, except that the MLD snooping feature snoops for IPv6 multicast traffic and operates on MLDv1 (RFC 2710) and MLDv2 (RFC 3810) control plane packets. MLD is a sub-protocol of Internet Control Message Protocol version 6 (ICMPv6), so MLD message types are a subset of ICMPv6 messages and MLD messages are identified in IPv6 packets by a preceding next header value of 58. Message types in MLDv1 include listener queries, multicast address-specific (MAS) queries, listener reports, and done messages. MLDv2 is designed to be interoperable with MLDv1 except that it has an extra query type, the multicast address and source-specific (MASS) query. The protocol level timers available in MLD are similar to those available in IGMP.

When MLD snooping is disabled, then all the multicast traffic is flooded to all the ports, whether they have an interest or not. When MLD snooping is enabled, the fabric will forward IPv6 multicast traffic based on MLD interest. Unknown IPv6 multicast traffic will be flooded based on the bridge-domain's IPv6 L3 unknown multicast flood setting.

There are two modes for forwarding unknown IPv6 multicast packets:

- Flooding mode: All endpoint groups (EPGs) and all ports under the bridge-domain will get the flooded packets.
- Optimized Multicast Flooding (OMF) mode: Only multicast router ports will get the packet.

Licensing Requirements for MLD

Product	License Requirement
NX-OS	MLD requires no license. Any feature not included in a license package is bundled with the nx-os image and is provided at no extra charge to you. For a complete explanation of the NX-OS licensing scheme, see the NX-OS Licensing Guide.

Prerequisites for MLD

MLD has the following prerequisites:

- You are logged into the device.
- For global configuration commands, you are in the correct virtual routing and forwarding (VRF) mode. The default configuration mode shown in the examples in this chapter applies to the default VRF.

Guidelines and Limitations for MLD

MLD has the following guidelines and limitations:

- Excluding or blocking a list of sources according to MLDv2 (RFC 3810) is not supported.
- When you modify the route-map to deny the multicast group, which is statically bound to the interface; the subsequent MLD reports are rejected by the local groups and the groups start ageing. The MLD leave message for the groups is allowed without any impact. This is a known and expected behaviour.
- MLD snooping is supported only on CN93240YC-FX2 with vPC and without vPC.==

Default Settings for MLD

Table 6: Default MLD Parameters

Parameters	Default
MLD version	2
Startup query interval	30 seconds
Startup query count	2
Robustness value	2
Querier timeout	255 seconds
Query timeout	255 seconds
Query max response time	10 seconds
Query interval	125 seconds
Last member query response interval	1 second

Parameters	Default
Last member query count	2
Group membership timeout	260 seconds
Report link local multicast groups	Disabled
Immediate leave	Disabled

Configuring MLD Parameters

You can configure the MLD global and interface parameters to affect the operation of the MLD process.

Note Before you can access the MLD commands, you must enable the MLD feature.

Configuring MLD Interface Parameters

Table 7: MLD Interface Parameters

Parameter	Description
MLD version	The MLD version that is enabled on the interface. MLDv2 supports MLDv1. The MLD version can be 1 or 2. The default is 2.
Static multicast groups	Multicast groups that are statically bound to the interface. You can configure the groups to join the interface with the (*, G) state or specify a source IP to join with the (S, G) state. You can specify a route-map policy name that lists the group prefixes, group ranges, and source prefixes to use with the **match ip multicast** command. **Note** Although you can configure the (S, G) state, the source tree is built only if you enable MLDv2. You can configure a multicast group on all the multicast-capable routers on the network so that pinging the group causes all the routers to respond.

Parameter	Description
Static multicast groups on OIF	Multicast groups that are statically bound to the output interface. You can configure the groups to join the output interface with the (*, G) state or specify a source IP to join with the (S, G) state. You can specify a route-map policy name that lists the group prefixes, group ranges, and source prefixes to use with the **match ip multicast** command. Although you can configure the (S, G) state, the source tree is built only if you enable MLDv2. **Note** Group prefixes in the route map must have a mask of 120 or longer.
Startup query interval	Startup query interval. By default, this interval is shorter than the query interval so that the software can establish the group state as quickly as possible. Values range from 1 to 18,000 seconds. The default is 30 seconds.
Startup query count	The number of queries sent at startup that are separated by the startup query interval. Values range from 1 to 10. The default is 2.
Robustness value	A robustness variable that you can tune to reflect expected packet loss on a congested network. You can increase the robustness variable to increase the number of times that packets are resent. Values range from 1 to 7. The default is 2.
Querier timeout	The number of seconds that the software waits after the previous querier has stopped querying and before it takes over as the querier. Values range from 1 to 65,535 seconds. The default is 255 seconds.
Query max response time	The maximum response time advertised in MLD queries. You can tune the burstiness of MLD messages on the network by setting a larger value so that host responses are spread out over a longer time. This value must be less than the query interval. Values range from 1 to 25 seconds. The default is 10 seconds.
Query interval	The frequency at which the software sends MLD host query messages. You can tune the number of MLD messages on the network by setting a larger value so that the software sends MLD queries less often. Values range from 1 to 18,000 seconds. The default is 125 seconds.

Parameter	Description
Last member query response interval	The query interval for response to an MLD query that the software sends after receiving a host leave message from the last known active host on the subnet. If no reports are received in the interval, the group state is deleted. You can use this value to tune how quickly the software stops transmitting on the subnet. The software can detect the loss of the last member of a group or source more quickly when the values are smaller. Values range from 1 to 25 seconds. The default is 1 second.
Last member query count	The number of times that the software sends an MLD query, separated by the last member query response interval, in response to a host leave message from the last known active host on the subnet. Values range from 1 to 5. The default is 2. **Caution** Setting this value to 1 means that a missed packet in either direction causes the software to remove the multicast state from the queried group or channel. The software can wait until the next query interval before the group is added again.
Group membership timeout	The group membership interval that must pass before the router decides that no members of a group or source exist on the network. Values range from 3 to 65,535 seconds. The default is 260 seconds.
Report link local multicast groups	An option that enables sending reports for groups in FF02::0/16. Link local addresses are used only by protocols on the local network. Reports are always sent for nonlink local groups. The default is disabled.
Report policy	An access policy for MLD reports that is based on a route-map policy.
Access groups	An option that configures a route-map policy to control the multicast groups that hosts on the subnet serviced by an interface can join. **Note** Only the **match ip multicast group** command is supported in this route map policy. The **match ip address** command for matching an ACL is not supported.

Parameter	Description
Immediate leave	An option that minimizes the leave latency of MLDv1 group memberships on a given MLD interface because the device does not send group-specific queries. When immediate leave is enabled, the device will remove the group entry from the multicast routing table immediately upon receiving a leave message for the group. The default is disabled. **Note** Use this command only when there is one receiver behind the interface for a given group.

[2] To configure route-map policies, see the CN93240YC-FX2 *NX-OS Unicast Routing Configuration Guide*.

Procedure

	Command or Action	Purpose

Step 1	**configure terminal**
	Example:
	`switch# configure terminal`
	`switch(config)#`
		Enters global configuration mode.

Step 2	**interface** *interface*
	Example:
	`switch(config)# interface ethernet 2/1`
	`switch(config-if)#`
		Enters interface configuration mode.

Step 3	**Option**: **ipv6 mld version** *value* Example `switch(config-if)# ipv6 mld version 2`	**Description**: Sets the MLD version that is enabled on the interface. MLDv2 supports MLDv1. Values can be 1 or 2. The default is 2. The *no* form of the command sets the version to 2.
	ipv6 mld join-group {**group** [**source** *source*] | **route-map** *policy-name*} Example `switch(config-if)# ipv6 mld join-group FFFE::1`	Statically binds a multicast group to the interface. If you specify only the group address, the (*, G) state is created. If you specify the source address, the (S, G) state is created. You can specify

Command or Action	Purpose
Option	**Description**
	a route-map policy name that lists the group prefixes, group ranges, and source prefixes to use with the **match ip multicast** command. **Note** A source tree is built for the (S, G) state only if you enable MLDv2. **Caution** The device CPU must handle the traffic generated by using this command.
ipv6 mld static-oif {**group** [**source** *source*] | **route-map** *policy-name*} Example `switch(config-if)# ipv6 mld static-oif FFFE::1`	Statically binds a multicast group to the outgoing interface, which is handled by the device hardware. If you specify only the group address, the (*, G) state is created. If you specify the source address, the (S, G) state is created. You can specify a route-map policy name that lists the group prefixes, group ranges, and source prefixes to use with the **match ip multicast** command. **Note** A source tree is built for the (S, G) state only if you enable MLDv2.

Command or Action	Purpose
Option	**Description**
	Note The maximum number of groups supported per entry in the route map is 256.
ipv6 mld startup-query-interval *seconds* Example `switch(config-if)# ipv6 mld startup-query-interval 25`	Sets the query interval used when the software starts up. Values can range from 1 to 18,000 seconds. The default is 31 seconds.
ipv6 mld startup-query-count *count* Example `switch(config-if)# ipv6 mld startup-query-count 3`	Sets the query count used when the software starts up. Values can range from 1 to 10. The default is 2.
ipv6 mld robustness-variable *value* Example `switch(config-if)# ipv6 mld robustness-variable 3`	Sets the robustness variable. You can use a larger value for a network prone to packet loss. Values can range from 1 to 7. The default is 2.
ipv6 mld querier-timeout *seconds* Example `switch(config-if)# ipv6 mld querier-timeout 300`	Sets the querier timeout that the software uses when deciding to take over as the querier. Values can range from 1 to 65,535 seconds. The default is 255 seconds.
ipv6 mld query-timeout *seconds* Example `switch(config-if)# ipv6 mld query-timeout 300`	Sets the query timeout that the software uses when deciding to take over as the querier. Values can range from 1 to 65,535 seconds. The default is 255 seconds.

Command or Action	Purpose
Option	**Description**
	Note This command has the same functionality as the **ipv6 mld querier-timeout** command.
ipv6 mld query-max-response-time *seconds* Example `switch(config-if)# ipv6 mld query-max-response-time 15`	Sets the response time advertised in MLD queries. Values can range from 1 to 25 seconds. The default is 10 seconds.
ipv6 mld query-interval *interval* Example `switch(config-if)# ipv6 mld query-interval 100`	Sets the frequency at which the software sends MLD host query messages. Values can range from 1 to 18,000 seconds. The default is 125 seconds.
ipv6 mld last-member-query-response-time *seconds* Example `switch(config-if)# ipv6 mld last-member-query-response-time 3`	Sets the query response time after sending membership reports before the software deletes the group state. Values can range from 1 to 25 seconds. The default is 1 second.
ipv6 mld last-member-query-count *count* Example `switch(config-if)# ipv6 mld last-member-query-count 3`	Sets the number of times that the software sends an MLD query in response to a host leave message. Values can range from 1 to 5. The default is 2.
ipv6 mld group-timeout seconds Example `switch(config-if)# ipv6 mld group-timeout 300`	Sets the group membership timeout for MLDv2. Values can range from 3 to 65,535 seconds. The default is 260 seconds.
ipv6 mld report-link-local-groups Example	Enables sending reports for groups in 224.0.0.0/24. Reports are always sent for nonlink

Command or Action	Purpose
Option	**Description**
`switch(config-if)# ipv6 mld report-link-local-groups`	local groups. By default, reports are not sent for link local groups.
ipv6 mld report-policy *policy* Example `switch(config-if)# ipv6 mld report-policy my_report_policy`	Configures an access policy for MLD reports that is based on a route-map policy.
ipv6 mld access-group *policy* Example `switch(config-if)# ipv6 mld access-group my_access_policy`	Configures a route-map policy to control the multicast groups that hosts on the subnet serviced by an interface can join. **Note** Only the **match ip multicast group** command is supported in this route map policy. The **match ip address** command for matching an ACL is not supported.
ipv6 mld immediate-leave Example `switch(config-if)# ipv6 mld immediate-leave`	Enables the device to remove the group entry from the multicast routing table immediately upon receiving a leave message for the group. Use this command to mnimize the leave latency of MLDv1 group memberships on a given MLD interface because the device does not send group-specific queries. The default is disabled.

	Command or Action	Purpose
	Option	**Description**
		Note Use this command only when there is one receiver behind the interface for a given group.
Step 4	(Optional) **show ipv6 mld interface** [*interface*] [**vrf** *vrf-name* | **all**] [**brief**] **Example:** `switch(config)# show ipv6 mld interface`	Displays MLD information about the interface.
Step 5	(Optional) **copy running-config startup-config** **Example:** `switch(config)# copy running-config startup-config`	Copies the running configuration to the startup configuration.

Configuring an MLD SSM Translation

You can configure an SSM translation to provide SSM support when the router receives MLDv1 listener reports. Only MLDv2 provides the capability to specify group and source addresses in listener reports. By default, the group prefix range is FF3x/96.

Table 8: Example SSM Translations

Group Prefix	Source Address
FF30::0/16	2001:0DB8:0:ABCD::1
FF30::0/16	2001:0DB8:0:ABCD::2
FF30:30::0/24	2001:0DB8:0:ABCD::3
FF32:40::0/24	2001:0DB8:0:ABCD::4

The following table shows the resulting M6RIB routes that the MLD process creates when it applies an SSM translation to the MLD v1 listener report. If more than one translation applies, the router creates the (S, G) state for each translation.

Table 9: Example Result of Applying SSM Translations

MLDv1 Listener Report	Resulting M6RIB Route
FF32:40::40	(2001:0DB8:0:ABCD::4, FF32:40::40)

MLDv1 Listener Report	Resulting M6RIB Route
FF30:10::10	(2001:0DB8:0:ABCD::1, FF30:10::10) (2001:0DB8:0:ABCD::2, FF30:10::10)

Procedure

	Command or Action	Purpose

Step 1	**configure terminal**
	Example:
	`switch# configure terminal`
	`switch(config)#`
		Enters global configuration mode.

Step 2	**ipv6** [icmp] **mld ssm-translate** *group-prefix source-addr* **Example:** `switch(config)# ipv6 mld ssm-translate FF30::0/16 2001:0DB8:0:ABCD::1`	Configures the translation of MLDv1 listener reports by the MLD process to create the (S, G) state as if the router had received an MLDv2 listener report.
Step 3	(Optional) **show running-configuration ssm-translate** **Example:** `switch(config)# show running-configuration ssm-translate`	Shows *ssm-translate* configuration lines in the running configuration.
Step 4	(Optional) **copy running-config startup-config** **Example:** `switch(config)# copy running-config startup-config`	Copies the running configuration to the startup configuration.

Verifying the MLD Configuration

To display the MLD configuration information, perform one of the following tasks:

show ipv6 mld interface [*interface*] [**vrf** *vrf-name* | **all**] [**brief**]	Displays MLD information about all interfaces or a selected interface or about the default VRF, a selected VRF, or all VRFs.
show ipv6 mld groups [*group* | *interface*] [**vrf** *vrf-name* | **all**]	Displays the MLD attached group membership for a group or interface or for the default VRF, a selected VRF, or all VRFs.
show ipv6 mld route [*group* | *interface*] [**vrf** *vrf-name* | **all**]	Displays the MLD attached group membership for a group or interface or for the default VRF, a selected VRF, or all VRFs.

show ipv6 mld local-groups	Displays the MLD local group membership.

Configuring MLD Snooping

MLD snooping can be enabled and disabled in the global configuration mode as well as in the VLAN configuration mode. Snooping is disabled by default in the global configuration mode and enabled per VLAN. Snooping is operational on a VLAN only if it is enabled both on the VLAN as well is in the global configuration mode.

Procedure

	Command or Action	Purpose

Step 1	**configure terminal**
	Example:
	`switch# configure terminal`
	`switch(config)#`
		Enters global configuration mode.

Step 2	**ipv6 mld snooping** **Example:** `switch(config)# ipv6 mld snooping`	Enables the admin state of the MLD snooping.
Step 3	**system mld snooping** **Example:** `switch(config)# system mld snooping`	This is an additional requirement to enable the MLD snooping on the CN93240YC-FX2 platform. Both step 2 and step 3 are required to completely enable snooping on the CN93240YC-FX2 platform.
Step 4	**hardware access-list tcam region** *ing-sup tcam-size* **Example:** `switch(config)# hardware access-list tcam region ing-sup 768`	Configures the TCAM region ing-sup to be 768 or more. **Note** After performing steps 3 and 4, you will be prompted to save the configuration and reboot the system for carving out the ACL and enable different hardware programming for v6 and v4 routerg.
Step 5	**ipv6 mld snooping explicit-tracking** **Example:** `switch(config)# ipv6 mld snooping explicit-tracking`	Enables or disables Explicit Host Tracking on a per VLAN basis. This command is enabled by default for both the MLD versions (v1 and v2).
Step 6	**ipv6 mld snooping report-suppression** **Example:** `switch(config)# ipv6 mld snooping report-suppression`	Enables or disables the report suppression. Every MLDv1 membership report received from the host is forwarded to all multicast router ports. When the report suppression is disabled, proxy reporting does not happen as

	Command or Action	Purpose
		all the MLD membership reports are forwarded to the router as is. This command is enabled by default.
Step 7	**ipv6 mld snooping v2-report-suppression** **Example:** `switch(config)# ipv6 mld snooping v2-report-suppression`	Enables MLDv2 report suppression. MLDv2 report suppression is disabled by default.
Step 8	**ipv6 mld snooping link-local-groups-suppression** **Example:** `switch(config)# ipv6 mld snooping link-local-groups-suppression`	Configures link-local-groups-suppression.
Step 9	**ipv6 mld snooping event-history vlan size** {**disabled** |**large** |**medium** |**small**} **Example:** `switch(config)# ipv6 mld snooping event-history vlan size medium`	Configures event history buffers for VLANs. Default value is medium.
Step 10	**ipv6 mld snooping event-history vlan-events** {**disabled** |**large** |**medium** |**small**} **Example:** `switch(config)# ipv6 mld snooping event-history vlan-events medium`	Configures event history buffers for VLAN events. Default value is medium.
Step 11	**ipv6 mld snooping event-history MLD-snoop-internal size** {**disabled** |**large** |**medium** |**small**} **Example:** `switch(config)# ipv6 mld snooping event-history MLD-snoop-internal size small`	Configures event history buffers for MLD-snoop internal events. Default value is small.
Step 12	**ipv6 mld snooping event-history mfdm size** {**disabled** |**large** |**medium** |**small**} **Example:** `switch(config)# ipv6 mld snooping event-history mfdm size small`	Configures event history buffers for MLD-snoop MFDM events. Default value is small.
Step 13	**ipv6 mld snooping event-history mfdm-sum** {**disabled** |**large** |**medium** |**small**} **Example:** `switch(config)# ipv6 mld snooping event-history mfdm-sum size small`	Configures event history buffers for MLD-snoop MFDM event summary. Default value is small.

	Command or Action	Purpose
Step 14	**ipv6 mld snooping event-history vpc size** {**disabled** |**large** |**medium** |**small**} **Example:** `switch(config)# ipv6 mld snooping event-history vpc size small`	Configures event history buffers for MLD-snoop vPC events. Default value is small.
Step 15	**vlan configuration** *vlan-id* **Example:** `switch(config)# vlan configuration 6`	Enters VLAN configuration mode.
Step 16	**[no] ipv6 mld snooping** **Example:** `switch(config-vlan)# no ipv6 mld snooping`	Disables or enables MLD snooping per VLAN. Once disabled, PIM6 will not work on the corresponding “interface vlan”.
Step 17	**ipv6 mld snooping fast-leave** **Example:** `switch(config-vlan)# ipv6 mld snooping fast-leave`	Allows you to turn on or off the fast-leave feature on a per-VLAN basis. This applies to MLDv2 hosts and is used on ports that are known to have only one host doing MLD behind that port. This command is disabled by default. This is a VLAN mode command.
Step 18	**ipv6 mld snooping mrouter interface** *interface-identifier* **Example:** `switch(config-vlan)# ipv6 mld snooping mrouter interface port-channel 1`	Specifies a static connection to a multicast router. The interface to the router must be in the VLAN where the command is entered and must be administratively up along with the line protocol. This is a VLAN mode command.
Step 19	**ipv6 mld snooping static-group** *group* [**source** *source*] **interface** *interface-identifier* **Example:** `switch(config-vlan)# ipv6 mld snooping static-group ff1e::abcd interface port-channel 2`	Configures a Layer2 port on a specific VLAN as a member of a multicast group statically. This is a VLAN mode command.
Step 20	**ipv6 mld snooping last-member-query-interval** [*interval*] **Example:** `switch(config-vlan)# ipv6 mld snooping last-member-query-interval 9`	Configures the interval for which the switch waits after sending a group-specific query to determine if hosts are still interested in a specific multicast group. It configures the interval for the MLD queries sent by the switch. Default is 1 second. Valid range is 1 to 25 seconds. This is a VLAN mode command. When both MLD fast-leave processing and the MLD query interval are configured, fast-leave processing is considered as the priority.

	Command or Action	Purpose
Step 21	**ipv6 mld snooping querier** *link-local address* **Example:** `switch(config-vlan)# ipv6 mld snooping querier aaaa::abcd`	Enables or disables IPv6 MLD snooping querier processing. MLD snooping querier supports the MLD snooping in a VLAN where PIM and MLD are not configured because the multicast traffic does not need to be routed.

Verifying the MLD Snooping Configuration

To display the MLD snooping configuration information, perform one of the following tasks:

show ipv6 mld interface [*interface*] [**vrf** *vrf-name* | **all**] [**brief**]	Displays MLD information about all interfaces or a selected interface or about the default VRF, a selected VRF, or all VRFs.
show ipv6 mld snooping [**vlan** *vlan-id*]	Displays the MLD snooping status and details for a given VLAN or all VLANs.
show ipv6 mld snooping mrouter [**vlan** *vlan-id*]	Displays the multicast router ports in each VLAN.
show ipv6 mld snooping querier [**vlan** *vlan-id*]	Displays details on the MLD Querier for the VLAN in which MLD Snooping is enabled.
show ipv6 mld snooping explicit-tracking vlan *vlan-id*	Displays the MLD snooping explicit tracking information.
show ipv6 mld snooping statistics global	Displays the global MLD snooping statistics.
show ipv6 mld snooping groups [**vlan** *vlan-id*] [**detail**]	Displays groups, the type of reports that are received for the group (host type) and the list of ports on which reports are received. The list of ports does not include the multicast router ports. This represents the list of ports on which the reports have been received and not the complete forwarding port set for the group. Displays the router ports by the */* entry in the non-detailed output.

Configuration Example for MLD

The following example shows how to configure MLD:

```
configure terminal
  ipv6 mld ssm-translate FF30::0/16 2001:0DB8:0:ABCD::1
  interface ethernet 2/1
    ipv6 mld version 2
    ipv6 mld join-group FFFE::1
    ipv6 mld startup-query-interval 25
    ipv6 mld startup-query-count 3
    ipv6 mld robustness-variable 3
    ipv6 mld querier-timeout 300
    ipv6 mld query-timeout 300
    ipv6 mld query-max-response-time 15
    ipv6 mld query-interval 100
    ipv6 mld last-member-query-response-time 3
    ipv6 mld last-member-query-count 3
    ipv6 mld group-timeout 300
    ipv6 mld report-link-local-groups
    ipv6 mld report-policy my_report_policy
    ipv6 mld access-group my_access_policy
```

CHAPTER 6

Configuring IGMP Snooping

This chapter describes how to configure Internet Group Management Protocol (IGMP) snooping on a NX-OS device.

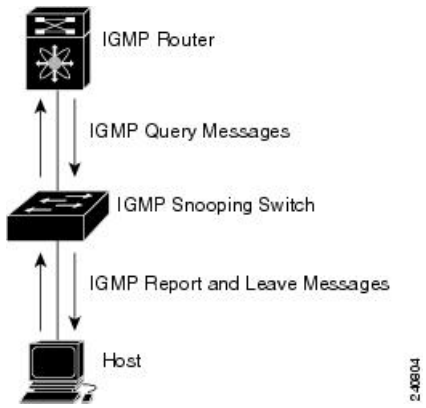
About IGMP Snooping

Note We recommend that you do not disable IGMP snooping on the device. If you disable IGMP snooping, you might see reduced multicast performance because of excessive false flooding within the device.

IGMP snooping software examines Layer 2 IP multicast traffic within a VLAN to discover the ports where interested receivers reside. Using the port information, IGMP snooping can reduce bandwidth consumption in a multi-access LAN environment to avoid flooding the entire VLAN. IGMP snooping tracks which ports are attached to multicast-capable routers to help the routers forward IGMP membership reports. The IGMP snooping software responds to topology change notifications. By default, IGMP snooping is enabled on the device.

This figure shows an IGMP snooping switch that sits between the host and the IGMP router. The IGMP snooping switch snoops the IGMP membership reports and Leave messages and forwards them only when necessary to the connected IGMP routers.

Figure 15: IGMP Snooping Switch



The IGMP snooping software operates upon IGMPv1, IGMPv2, and IGMPv3 control plane packets where Layer 3 control plane packets are intercepted and influence the Layer 2 forwarding behavior.

The NX-OS IGMP snooping software has the following proprietary features:

- Source filtering that allows forwarding of multicast packets based on destination and source IP addresses
- Multicast forwarding based on IP addresses rather than the MAC address
- Multicast forwarding alternately based on the MAC address

For more information about IGMP snooping, see RFC 4541.

IGMPv1 and IGMPv2

Both IGMPv1 and IGMPv2 support membership report suppression, which means that if two hosts on the same subnet want to receive multicast data for the same group, the host that receives a member report from the other host suppresses sending its report. Membership report suppression occurs for hosts that share a port.

If no more than one host is attached to each VLAN switch port, you can configure the fast leave feature in IGMPv2. The fast leave feature does not send last member query messages to hosts. As soon as the software receives an IGMP leave message, the software stops forwarding multicast data to that port.

IGMPv1 does not provide an explicit IGMP leave message, so the software must rely on the membership message timeout to indicate that no hosts remain that want to receive multicast data for a particular group.

Note The software ignores the configuration of the last member query interval when you enable the fast leave feature because it does not check for remaining hosts.

IGMPv3

The IGMPv3 snooping implementation on NX-OS supports full IGMPv3 snooping, which provides constrained flooding based on the (S, G) information in the IGMPv3 reports. This source-based filtering enables the device to constrain multicast traffic to a set of ports based on the source that sends traffic to the multicast group.

By default, the software tracks hosts on each VLAN port. The explicit tracking feature provides a fast leave mechanism. Because every IGMPv3 host sends membership reports, report suppression limits the amount of traffic that the device sends to other multicast-capable routers. When report suppression is enabled, and no IGMPv1 or IGMPv2 hosts requested the same group, the software provides proxy reporting. The proxy feature builds the group state from membership reports from the downstream hosts and generates membership reports in response to queries from upstream queriers.

Even though the IGMPv3 membership reports provide a full accounting of group members on a LAN segment, when the last host leaves, the software sends a membership query. You can configure the parameter last member query interval. If no host responds before the timeout, the software removes the group state.

IGMP Snooping Querier

When PIM is not enabled on an interface because the multicast traffic does not need to be routed, you must configure an IGMP snooping querier to send membership queries. You define the querier in a VLAN that contains multicast sources and receivers but no other active querier.

The querier can be configured to use any IP address in the VLAN.

As a best practice, a unique IP address, one that is not already used by the switch interface or the Hot Standby Router Protocol (HSRP) virtual IP address, should be configured so as to easily reference the querier.

Note The IP address for the querier should not be a broadcast IP address, multicast IP address, or 0 (0.0.0.0).

When an IGMP snooping querier is enabled, it sends out periodic IGMP queries that trigger IGMP report messages from hosts that want to receive IP multicast traffic. IGMP snooping listens to these IGMP reports to establish appropriate forwarding.

The IGMP snooping querier performs querier election as described in RFC 2236. Querier election occurs in the following configurations:

- When there are multiple switch queriers configured with the same subnet on the same VLAN on different switches.
- When the configured switch querier is in the same subnet as with other Layer 3 SVI queriers.

Virtualization Support

You can define multiple virtual routing and forwarding (VRF) instances for IGMP snooping.

You can use the **show** commands with a VRF argument to provide a context for the information displayed. The default VRF is used if no VRF argument is supplied.

For information about configuring VRFs, see the *CN93240YC-FX2 NX-OS Unicast Routing Configuration Guide*.

Licensing Requirements for IGMP Snooping

Product	License Requirement
NX-OS	IGMP snooping requires no license. Any feature not included in a license package is bundled with the nx-os image and is provided at no extra charge to you. For a complete explanation of the NX-OS licensing scheme, see *NX-OS Licensing Guide*.

Prerequisites for IGMP Snooping

IGMP snooping has the following prerequisites:

- You are logged onto the device.
- For global commands, you are in the correct virtual routing and forwarding (VRF) mode. The default configuration mode shown in the examples in this chapter applies to the default VRF.

Guidelines and Limitations for IGMP Snooping

IGMP snooping has the following guidelines and limitations:

- CN93240YC-FX2 switches support IGMP snooping for IPv4 but do not support MLD snooping for IPv6.
- IGMP snooping is not supported with PVLAN.
- Layer 3 IPv6 multicast routing is not supported.
- Layer 2 IPv6 multicast packets will be flooded on the incoming VLAN.
- IGMP snooping configuration must be identical on both vPC peers in a vPC pair. Either enable or disable IGMP snooping on both vPC peers.=

Note Enabling or disabling IGMP snooping on both vPC peers also enables the forwarding of IGMP queries from different MVR source VLANs into the same MVR receiver VLAN. The resulting IGMP queries may send out queries with different versions and query interval. If you prefer to maintain the behavior prior to NX-OS Release 7.0(3)I3(1) use the **mvr-suppress-query vlan** <*id*> command.

- In releases prior to NX-OS Release 7.0(3)I3(1) if you are configuring vPC peers, the differences in the IGMP snooping configuration options between the two devices have the following results:=

- If IGMP snooping is enabled on one device but not on the other, the device on which snooping is disabled floods all multicast traffic.
- A difference in multicast router or static group configuration can cause traffic loss.
- The fast leave, explicit tracking, and report suppression options can differ if they are used for forwarding traffic.
- If a query parameter is different between the devices, one device expires the multicast state faster while the other device continues to forward. This difference results in either traffic loss or forwarding for an extended period.
- If an IGMP snooping querier is configured on both devices, only one of them will be active because an IGMP snooping querier shuts down if a query is seen in the traffic.

- You must enable the **ip igmp snooping group-timeout** command when you use the **ip igmp snooping proxy general-queries** command. We recommend that you set it to "never". Otherwise, you might experience multicast packet loss.
- All external multicast router ports (either statically configured or dynamically learned) use the global ltl index. As a result, traffic in VLAN X goes out on the multicast router ports in both VLAN X and VLAN Y, in case both multicast router ports (Layer 2 trunks) carry both VLAN X and VLAN Y.
- When you modify the route-map to deny the multicast group, which is statically bound to the interface; the subsequent IGMP reports are rejected by the local groups and the groups start ageing. The IGMP leave message for the groups is allowed without any impact. This is a known and expected behaviour.

Default Settings

Parameters	Default
IGMP snooping	Enabled
Explicit tracking	Enabled
Fast leave	Disabled
Last member query interval	1 second
Snooping querier	Disabled
Report suppression	Enabled
Link-local groups suppression	Enabled
Optimise-multicast-flood	Disabled
IGMPv3 report suppression for the entire device	Disabled
IGMPv3 report suppression per VLAN	Enabled

Configuring IGMP Snooping Parameters

Note If you are familiar with the IOS CLI, be aware that the NX-OS commands for this feature might differ from the IOS commands that you would use.

Note You must enable IGMP snooping globally before any other commands take effect.

Configuring Global IGMP Snooping Parameters

To affect the operation of the IGMP snooping process globally, you can configure various optional IGMP snooping parameters.

Notes for IGMP Snooping Parameters

- IGMP Snooping Proxy parameter

  To decrease the burden placed on the snooping switch during each IGMP general query (GQ) interval, the NX-OS software provides a way to decouple the periodic general query behavior of the IGMP snooping switch from the query interval configured on the multicast routers.

  You can configure the device to consume IGMP general queries from the multicast router, rather than flooding the general queries to all the switchports. When the device receives a general query, it produces proxy reports for all currently active groups and distributes the proxy reports over the period specified by the MRT that is specified in the router query. At the same time, independent of the periodic general query activity of the multicast router, the device sends an IGMP general query on each port in the VLAN in a round-robin fashion. It cycles through all the interfaces in the VLAN at the rate given by the following formula.
  Rate = {number of interfaces in VLAN} * {configured MRT} * {number of VLANs}

  When queries are run in this mode, the default MRT value is 5,000 milliseconds (5 seconds). For a device that has 500 switchports in a VLAN, it would take 2,500 seconds (40 minutes) to cycle through all the interfaces in the system. This is also true when the device itself is the querier.

  This behavior ensures that only one host responds to a general query at a given time, and it keeps the simultaneous reporting rate below the packet-per-second IGMP capability of the device (approximately 3,000 to 4,000 pps).=

  Note When you use this option, you must change the **ip igmp snooping group-timeout** parameter to a high value or to never time out.

  The **ip igmp snooping proxy general-queries** [**mrt**] command causes the snooping function to proxy reply to general queries from the multicast router while also sending round-robin general queries on each switchport with the specified MRT value. (The default MRT value is 5 seconds.)

- IGMP Snooping Group-timeout parameter

Configuring the group-timeout parameter disables the behavior of an expiring membership based on three missed general queries. Group membership remains on a given switchport until the device receives an explicit IGMP leave on that port.

The **ip igmp snooping group-timeout** {*timeout* | **never**} command modifies or disables the behavior of an expiring IGMP snooping group membership after three missed general queries.

Procedure

Step 1 **configure terminal**

Example:

```
switch# configure terminal
switch(config)#
```

Enters global configuration mode.

Step 2 Use the following commands to configure global IGMP snooping parameters.

Option	Description
ip igmp snooping `switch(config)# ip igmp snooping`	Enables IGMP snooping for the device. The default is enabled. **Note** If the global setting is disabled with the **no** form of this command, IGMP snooping on all VLANs is disabled, whether IGMP snooping is enabled on a VLAN or not. If you disable IGMP snooping, Layer 2 multicast frames flood to all modules.
ip igmp snooping event-history `switch(config)# ip igmp snooping event-history`	Configures the size of the event history buffer. The default is small.
ip igmp snooping group-timeout {*minutes* | **never**} `switch(config)# ip igmp snooping group-timeout never`	Configures the group membership timeout value for all VLANs on the device.
ip igmp snooping link-local-groups-suppression `switch(config)# ip igmp snooping link-local-groups-suppression`	Configures link-local groups suppression for the entire device. The default is enabled.
ip igmp snooping optimise-multicast-flood	Optimizes optimized multicast flooding (OMF) on all VLANs on the device. The default is disabled.

Option	Description
`switch(config)# ip igmp snooping optimise-multicast-flood`	
ip igmp snooping proxy general-inquiries [**mrt** *seconds*] `switch(config)# ip igmp snooping proxy general-inquiries`	Configures the IGMP snooping proxy for the device. The default is 5 seconds.
ip igmp snooping v3-report-suppression `switch(config)# ip igmp snooping v3-report-suppression`	Limits the membership report traffic sent to multicast-capable routers. When you disable report suppression, all IGMP reports are sent as-is to multicast-capable routers. The default is enabled.
ip igmp snooping report-suppression `switch(config)# ip igmp snooping report-suppression`	Configures IGMPv3 report suppression and proxy reporting. The default is disabled.

Step 3 **copy running-config startup-config**

Example:

```
switch(config)# copy running-config startup-config
```

(Optional) Copies the running configuration to the startup configuration.

Configuring IGMP Snooping Parameters per VLAN

To affect the operation of the IGMP snooping process per VLAN, you can configure various optional IGMP snooping parameters.

Note You configure the IGMP snooping parameters that you want by using this configuration mode; however, the configurations apply only after you specifically create the specified VLAN. See the *Cisco Nexus 9000 Series NX-OS Layer 2 Switching Configuration Guide* for information on creating VLANs.

Procedure

Step 1 **configure terminal**

Example:

```
switch# configure terminal
switch(config)#
```

Enters global configuration mode.

Step 2 **ip igmp snooping**

Example:

```
switch(config)# ip igmp snooping
```

Enables IGMP snooping. The default is enabled.

Note If the global setting is disabled with the **no** form of this command, IGMP snooping on all VLANs is disabled, whether IGMP snooping is enabled on a VLAN or not. If you disable IGMP snooping, Layer 2 multicast frames flood to all modules.

Step 3 **vlan configuration** *vlan-id*

Example:

```
switch(config)# vlan configuration 2
switch(config-vlan-config)#
```

Configures the IGMP snooping parameters you want for the VLAN. These configurations do not apply until you create the specified VLAN.

Step 4 Use the following commands to configure IGMP snooping parameters per VLAN.

Option	Description
ip igmp snooping `switch(config-vlan-config)# ip igmp snooping`	Enables IGMP snooping for the current VLAN. The default is enabled.
ip igmp snooping access-group {**prefix-list** | **route-map**} *policy-name* **interface** *interface slot/port* `switch(config-vlan-config)# ip igmp snooping access-group prefix-list plist interface ethernet 2/2`	Configures a filter for IGMP snooping reports that is based on a prefix-list or route-map policy. The default is disabled.
ip igmp snooping explicit-tracking `switch(config-vlan-config)# ip igmp snooping explicit-tracking`	Tracks IGMPv3 membership reports from individual hosts for each port on a per-VLAN basis. The default is enabled on all VLANs.
ip igmp snooping fast-leave	Supports IGMPv2 hosts that cannot be explicitly tracked because of the host report suppression mechanism of the

Option	Description
`switch(config-vlan-config)# ip igmp snooping fast-leave`	IGMPv2 protocol. When you enable fast leave, the IGMP software assumes that no more than one host is present on each VLAN port. The default is disabled for all VLANs.
ip igmp snooping group-timeout {*minutes* | **never**} `switch(config-vlan-config)# ip igmp snooping group-timeout never`	Configures the group membership timeout for the specified VLANs.
ip igmp snooping last-member-query-interval *seconds* `switch(config-vlan-config)# ip igmp snooping last-member-query-interval 3`	Removes the group from the associated VLAN port if no hosts respond to an IGMP query message before the last member query interval expires. Values range from 1 to 25 seconds. The default is 1 second.
ip igmp snooping proxy general-queries [**mrt** *seconds*] `switch(config-vlan-config)# ip igmp snooping proxy general-queries`	Configures an IGMP snooping proxy for specified VLANs. The default is 5 seconds.
[**no**] **ip igmp snooping proxy-leave use-group-address** `switch(config-vlan-config)# ip igmp snooping proxy-leave use-group-address`	Changes the destination address of proxy leave messages to the address of the group that is leaving. Normally, IGMP proxy leave messages generated by the IGMP snooping module use the 224.0.0.2 multicast router address when all hosts leave the group. You should implement this configuration if your multicast applications rely on receiving reports and leave messages to start or stop multicast traffic based on the destination address of the packet.
ip igmp snooping querier *ip-address* `switch(config-vlan-config)# ip igmp snooping querier 172.20.52.106`	Configures a snooping querier when you do not enable PIM because multicast traffic does not need to be routed. The IP address is used as the source in messages.
ip igmp snooping querier-timeout *seconds* `switch(config-vlan-config)# ip igmp snooping querier-timeout 300`	Configures a snooping querier timeout value for IGMPv2 when you do not enable PIM because multicast traffic does not need to be routed. The default is 255 seconds.
ip igmp snooping query-interval *seconds*	Configures a snooping query interval when you do not enable PIM because multicast traffic does not need to be routed. The default value is 125 seconds.

Option	Description
`switch(config-vlan-config)# ip igmp snooping query-interval 120`	
ip igmp snooping query-max-response-time *seconds* `switch(config-vlan-config)# ip igmp snooping query-max-response-time 12`	Configures a snooping MRT for query messages when you do not enable PIM because multicast traffic does not need to be routed. The default value is 10 seconds.

[**no**] **ip igmp snooping report-flood** {**all** | **interface ethernet** *slot/port*}
 	
`switch(config-vlan-config)# ip igmp snooping report-flood interface ethernet 1/2`
`ip igmp snooping report-flood interface ethernet 1/3`
	Floods IGMP reports on all active interfaces of the VLAN or only on specific interfaces.
	
	IGMP reports typically are forwarded to multicast router ports as detected by the IGMP snooping module and are not flooded in the VLAN. However, this command forces the switch to send IGMP reports to custom ports belonging to the VLAN in addition to the multicast router ports. You should implement this configuration if your multicast applications require the ability to view IGMP reports in order to transmit traffic.

ip igmp snooping report-policy {**prefix-list** | **route-map**} *policy-name* **interface** *interface slot/port* `switch(config-vlan-config)# ip igmp snooping report-policy route-map rmap interface ethernet 2/4`	Configures a filter for IGMP snooping reports that is based on a prefix-list or route-map policy. The default is disabled.
ip igmp snooping startup-query-count *value* `switch(config-vlan-config)# ip igmp snooping startup-query-count 5`	Configures snooping for a number of queries sent at startup when you do not enable PIM because multicast traffic does not need to be routed.
ip igmp snooping startup-query-interval *seconds* `switch(config-vlan-config)# ip igmp snooping startup-query-interval 15000`	Configures a snooping query interval at startup when you do not enable PIM because multicast traffic does not need to be routed.
ip igmp snooping robustness-variable *value* `switch(config-vlan-config)# ip igmp snooping robustness-variable 5`	Configures the robustness value for the specified VLANs. The default value is 2.

Option	Description
ip igmp snooping report-suppression `switch(config-vlan-config)# ip igmp snooping report-suppression`	Limits the membership report traffic sent to multicast-capable routers. When you disable report suppression, all IGMP reports are sent as-is to multicast-capable routers. The default is enabled.
ip igmp snooping mrouter interface *interface* `switch(config-vlan-config)# ip igmp snooping mrouter interface ethernet 2/1`	Configures a static connection to a multicast router. The interface to the router must be in the selected VLAN. You can specify the interface by the type and the number, such as **ethernet** *slot*/*port*.
ip igmp snooping static-group *group-ip-addr* [**source** *source-ip-addr*] **interface** *interface* `switch(config-vlan-config)# ip igmp snooping static-group 230.0.0.1 interface ethernet 2/1`	Configures the Layer 2 port of a VLAN as a static member of a multicast group. You can specify the interface by the type and the number, such as **ethernet** *slot*/*port*.
ip igmp snooping link-local-groups-suppression `switch(config-vlan-config)# ip igmp snooping link-local-groups-suppression`	Configures link-local groups suppression for the specified VLANs. The default is enabled.
ip igmp snooping v3-report-suppression `switch(config-vlan-config)# ip igmp snooping v3-report-suppression`	Configures IGMPv3 report suppression and proxy reporting for the specified VLANs. The default is enabled per VLAN.
ip igmp snooping version *value* `switch(config-vlan-config)# ip igmp snooping version 2`	Configures the IGMP version number for the specified VLANs.

Step 5 **copy running-config startup-config**

Example:

```
switch(config)# copy running-config startup-config
```

(Optional) Copies the running configuration to the startup configuration.

Verifying the IGMP Snooping Configuration

Command	Description
show ip igmp snooping [**vlan** *vlan-id*]	Displays the IGMP snooping configuration by VLAN.
show ip igmp snooping groups [*source* [*group*] | *group* [*source*]] [**vlan** *vlan-id*] [**detail**]	Displays IGMP snooping information about groups by VLAN.
show ip igmp snooping querier [**vlan** *vlan-id*]	Displays IGMP snooping queriers by VLAN.
show ip igmp snooping mroute [**vlan** *vlan-id*]	Displays multicast router ports by VLAN.
show ip igmp snooping explicit-tracking [**vlan** *vlan-id*] [**detail**]	Displays IGMP snooping explicit tracking information by VLAN. **Note** For vPC VLANs, you must enter the **detail** keyword to display this command on both vPC peer switches, beginning with Cisco NX-OS Release 7.0(3)I7(1). If you do not enter the **detail** keyword, this command displays only on the vPC switch that received the native report.

Displaying IGMP Snooping Statistics

You can display the IGMP snooping statistics using these commands.

Command	Description
show ip igmp snooping statistics vlan	Displays IGMP snooping statistics. You can see the virtual port channel (vPC) statistics in this output.
show ip igmp snooping {**report-policy** | **access-group**} **statistics** [**vlan** *vlan*]	Displays detailed statistics per VLAN when IGMP snooping filters are configured.

Clearing IGMP Snooping Statistics

You can clear the IGMP snooping statistics using these commands.

Command	Description
clear ip igmp snooping statistics vlan	Clears the IGMP snooping statistics.
clear ip igmp snooping {**report-policy** | **access-group**} **statistics** [**vlan** *vlan*]	Clears the IGMP snooping filter statistics.

Configuration Examples for IGMP Snooping

Note The configurations in this section apply only after you create the specified VLAN. See the *CN93240YC-FX2 NX-OS Layer 2 Switching Configuration Guide* for information on creating VLANs.

The following example shows how to configure the IGMP snooping parameters:

```
config t
  ip igmp snooping
  vlan configuration 2
    ip igmp snooping
    ip igmp snooping explicit-tracking
    ip igmp snooping fast-leave
    ip igmp snooping last-member-query-interval 3
    ip igmp snooping querier 172.20.52.106
    ip igmp snooping report-suppression
    ip igmp snooping mrouter interface ethernet 2/1
    ip igmp snooping static-group 230.0.0.1 interface ethernet 2/1
    ip igmp snooping link-local-groups-suppression
    ip igmp snooping v3-report-suppression
```

The following example shows how to configure prefix lists and use them to filter IGMP snooping reports:

```
ip prefix-list plist seq 5 permit 224.1.1.1/32
ip prefix-list plist seq 10 permit 224.1.1.2/32
ip prefix-list plist seq 15 deny 224.1.1.3/32
ip prefix-list plist seq 20 deny 225.0.0.0/8 eq 32

vlan configuration 2
  ip igmp snooping report-policy prefix-list plist interface Ethernet 2/2
  ip igmp snooping report-policy prefix-list plist interface Ethernet 2/3
```

In the above example, the prefix-list permits 224.1.1.1 and 224.1.1.2 but rejects 224.1.1.3 and all the groups in the 225.0.0.0/8 range. The prefix-list is an implicit "deny" if there is no match. If you wish to permit everything else, add **ip prefix-list plist seq 30 permit 224.0.0.0/4 eq 32**.

The following example shows how to configure route maps and use them to filter IGMP snooping reports:

```
route-map rmap permit 10
  match ip multicast group 224.1.1.1/32
route-map rmap permit 20
  match ip multicast group 224.1.1.2/32
route-map rmap deny 30
  match ip multicast group 224.1.1.3/32
route-map rmap deny 40
  match ip multicast group 225.0.0.0/8

vlan configuration 2
  ip igmp snooping report-policy route-map rmap interface Ethernet 2/4
  ip igmp snooping report-policy route-map rmap interface Ethernet 2/5
```

In the above example, the route-map permits 224.1.1.1 and 224.1.1.2 but rejects 224.1.1.3 and all the groups in the 225.0.0.0/8 range. The route-map is an implicit "deny" if there is no match. If you wish to permit everything else, add **route-map rmap permit 50 match ip multicast group 224.0.0.0/4**.

CHAPTER 8

Configuring MSDP

This chapter describes how to configure Multicast Source Discovery Protocol (MSDP) on a NX-OS device.

About MSDP

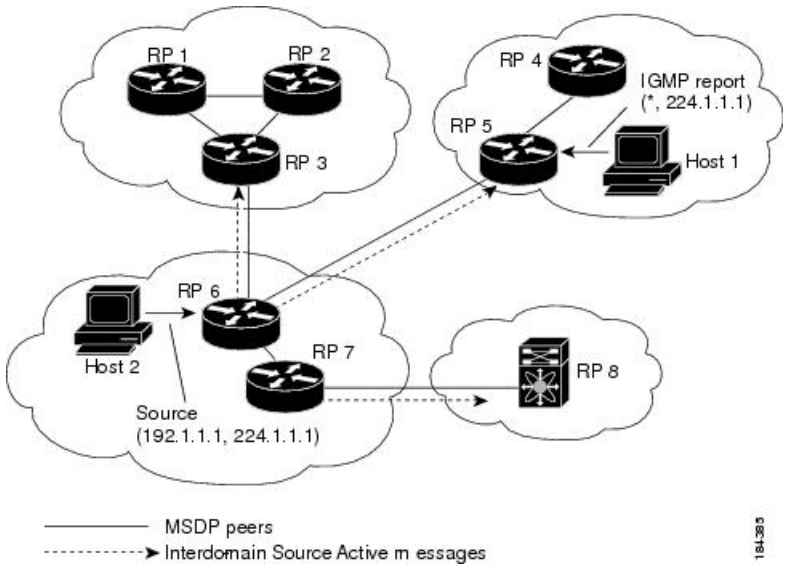
You can use the Multicast Source Discovery Protocol (MSDP) to exchange multicast source information between multiple Border Gateway Protocol (BGP) enabled Protocol Independent Multicast (PIM) sparse-mode domains. In addition, MSDP can be used to create an Anycast-RP configuration to provide RP redundancy and load sharing. For information about BGP, see the *CN93240YC-FX2 NX-OS Unicast Routing Configuration Guide*.

When a receiver joins a group that is transmitted by a source in another domain, the rendezvous point (RP) sends PIM join messages in the direction of the source to build a shortest path tree. The designated router (DR) sends packets on the sourcetree within the source domain, which can travel through the RP in the source domain and along the branches of the sourcetree to other domains. In domains where there are receivers, RPs in those domains can be on the sourcetree. The peering relationship is conducted over a TCP connection.

The following figure shows four PIM domains. The connected RPs (routers) are called MSDP peers because they are exchanging active source information with each other. Each MSDP peer advertises its own set of multicast source information to the other peers. Source Host 2 sends the multicast data to group 224.1.1.1. On RP 6, the MSDP process learns about the source through PIM register messages and generates Source-Active (SA) messages to its MSDP peers that contain information about the sources in its domain. When RP 3 and RP 5 receive the SA messages, they forward them to their MSDP peers. When RP 5 receives the request from

Host 1 for the multicast data on group 224.1.1.1, it builds a shortest path tree to the source by sending a PIM join message in the direction of Host 2 at 192.1.1.1.

Figure 16: MSDP Peering Between RPs in Different PIM Domains



When you configure MSDP peering between each RP, you create a full mesh. Full MSDP meshing is typically done within an autonomous system, as shown between RPs 1, 2, and 3, but not across autonomous systems. You use BGP to do a loop suppression and MSDP peer-RPF to suppress looping SA messages.

Note You do not need to configure BGP in order to use Anycast-RP (a set of RPs that can perform load balancing and failover) within a PIM domain.

Note You can use PIM Anycast (RFC 4610) to provide the Anycast-RP function instead of MSDP.

For detailed information about MSDP, see RFC 3618.

SA Messages and Caching

MSDP peers exchange Source-Active (SA) messages to propagate information about active sources. SA messages contain the following information:

- Source address of the data source
- Group address that the data source uses
- IP address of the RP or the configured originator ID

When a PIM register message advertises a new source, the MSDP process reencapsulates the message in an SA message that is immediately forwarded to all MSDP peers.

The SA cache holds the information for all sources learned through SA messages. Caching reduces the join latency for new receivers of a group because the information for all known groups can be found in the cache. You can limit the number of cached source entries by configuring the SA limit peer parameter. You can limit the number of cached source entries for a specific group prefix by configuring the group limit global parameter. The SA cache is enabled by default and cannot be disabled.

The MSDP software sends SA messages for each group in the SA cache every 60 seconds or at the configured SA interval global parameter. An entry in the SA cache is removed if an SA message for that source and group is not received within the SA interval plus 3 seconds.

MSDP Peer-RPF Forwarding

MSDP peers forward the SA messages that they receive away from the originating RP. This action is called peer-RPF flooding. The router examines the BGP or MBGP routing table to determine which peer is the next hop in the direction of the originating RP of the SA message. This peer is called a reverse path forwarding (RPF) peer.

If the MSDP peer receives the same SA message from a non-RPF peer in the direction of the originating RP, it drops the message. Otherwise, it forwards the message to all its MSDP peers.

MSDP Mesh Groups

You can use MSDP mesh groups to reduce the number of SA messages that are generated by peer-RPF flooding. By configuring a peering relationship between all the routers in a mesh and then configuring a mesh group of these routers, the SA messages that originate at a peer are sent by that peer to all other peers. SA messages received by peers in the mesh are not forwarded.

A router can participate in multiple mesh groups. By default, no mesh groups are configured.

Licensing Requirements for MSDP

Product	License Requirement
NX-OS	MSDP requires an Enterprise Services license. For a complete explanation of the NX-OS licensing scheme and how to obtain and apply licenses, see the *NX-OS Licensing Guide*.

Prerequisites for MSDP

MSDP has the following prerequisites:

- You are logged onto the device.
- For global commands, you are in the correct virtual routing and forwarding (VRF) mode. The default configuration mode shown in the examples in this chapter applies to the default VRF.
- You configured PIM for the networks where you want to configure MSDP.

Guidelines and Limitations for MSDP

MSDP has the following guidelines and limitations:

- Beginning with NX-OS Release 9.2(2) MSDP is supported on CN93240YC-FX2 platform switches.=

Default Settings

This table lists the default settings for MSDP parameters.

Table 17: Default MSDP Parameters

Parameters	Default
Description	Peer has no description
Administrative shutdown	Peer is enabled when it is defined
MD5 password	No MD5 password is enabled
SA policy IN	All SA messages are received
SA policy OUT	All registered sources are sent in SA messages
SA limit	No limit is defined
Originator interface name	RP address of the local system
Group limit	No group limit is defined
SA interval	60 seconds

Configuring MSDP

You can establish MSDP peering by configuring the MSDP peers within each PIM domain as follows:

1. Select the routers to act as MSDP peers.
2. Enable the MSDP feature.
3. Configure the MSDP peers for each router identified in Step 1.
4. Configure the optional MSDP peer parameters for each MSDP peer.
5. Configure the optional global parameters for each MSDP peer.
6. Configure the optional mesh groups for each MSDP peer.

Note The MSDP commands that you enter before you enable MSDP are cached and then run when MSDP is enabled. Use the **ip msdp peer** or **ip msdp originator-id** command to enable MSDP.

Note If you are familiar with the IOS CLI, be aware that the NX-OS commands for this feature might differ from the IOS commands that you would use.

Enabling the MSDP Feature

Procedure

	Command or Action	Purpose

Step 1	**configure terminal**
	Example:
	`switch# configure terminal`
	`switch(config)#`
		Enters global configuration mode.

Step 2	**feature msdp** **Example:** `switch# feature msdp`	Enables the MSDP feature so that you can enter MSDP commands. By default, the MSDP feature is disabled.
Step 3	(Optional) **show running-configuration msdp** **Example:** `switch# show running-configuration msdp`	Shows the running-configuration information for MSDP.
Step 4	(Optional) **copy running-config startup-config** **Example:** `switch(config)# copy running-config startup-config`	Copies the running configuration to the startup configuration.

Configuring MSDP Peers

You can configure an MSDP peer when you configure a peering relationship with each MSDP peer that resides either within the current PIM domain or in another PIM domain. MSDP is enabled on the router when you configure the first MSDP peering relationship.

Before you begin

Ensure that you have installed the Enterprise Services license and enabled PIM and MSDP.

Ensure that you configured PIM in the domains of the routers that you will configure as MSDP peers.

Procedure

	Command or Action	Purpose

Step 1	**configure terminal**
	Example:
	`switch# configure terminal`
	`switch(config)#`
		Enters global configuration mode.

Step 2	**ip msdp peer** *peer-ip-address* **connect-source** *interface* [**remote-as** *as-number*] **Example:** `switch(config)# ip msdp peer 192.168.1.10 connect-source ethernet 2/1 remote-as 8`	Configures an MSDP peer with the specified peer IP address. The software uses the source IP address of the interface for the TCP connection with the peer. The interface can take the form of *type slot/port*. If the AS number is the same as the local AS, then the peer is within the PIM domain; otherwise, this peer is external to the PIM domain. By default, MSDP peering is disabled. **Note** MSDP peering is enabled when you use this command.
Step 3	Repeat Step 2 for each MSDP peering relationship by changing the peer IP address, the interface, and the AS number as appropriate.	—
Step 4	(Optional) **show ip msdp summary** [**vrf** [*vrf-name* | **all**]] **Example:** `switch# show ip msdp summary`	Displays a summary of MDSP peers.
Step 5	(Optional) **copy running-config startup-config** **Example:** `switch(config)# copy running-config startup-config`	Copies the running configuration to the startup configuration.

Configuring MSDP Peer Parameters

You can configure the optional MSDP peer parameters described in this table. You configure these parameters in global configuration mode for each peer based on its IP address.

Table 18: MSDP Peer Parameters

Parameter	Description
Description	Description string for the peer. By default, the peer has no description.

Parameter	Description
Administrative shutdown	Method to shut down the MSDP peer. The configuration settings are not affected by this command. You can use this parameter to allow configuration of multiple parameters to occur before making the peer active. The TCP connection with other peers is terminated by the shutdown. By default, a peer is enabled when it is defined.
MD5 password	MD5-shared password key used for authenticating the peer. By default, no MD5 password is enabled.
SA policy IN	Route-map policy for incoming SA messages. By default, all SA messages are received. **Note** To configure route-map policies, see the *CN93240YC-FX2 NX-OS Unicast Routing Configuration Guide*.
SA policy OUT	Route-map policy for outgoing SA messages. By default, all registered sources are sent in SA messages. **Note** To configure route-map policies, see the *CN93240YC-FX2 NX-OS Unicast Routing Configuration Guide*.
SA limit	Number of (S, G) entries accepted from the peer and stored in the SA cache. By default, there is no limit.

Before you begin

Ensure that you have installed the Enterprise Services license and enabled PIM and MSDP.

Procedure

<table>
<tr><th></th><th>Command or Action</th><th>Purpose</th></tr>
<tr><td>Step 1</td><td>configure terminal
Example:
<pre>switch# configure terminal
switch(config)#</pre></td><td>Enters global configuration mode.</td></tr>
<tr><td>Step 2</td><td><table><tr><th>Option</th><th>Description</th></tr><tr><td>ip msdp description peer-ip-address description
Example:
<pre>switch(config)# ip msdp description</pre></td><td>Sets a description string for the peer. By default, the peer has no description.</td></tr></table></td><td>The following commands configure the MSDP peer parameters.</td></tr>
</table>

Command or Action		Purpose
Option	**Description**	
`192.168.1.10 peer in Engineering network`		
ip msdp shutdown *peer-ip-address* **Example:** `switch(config)# ip msdp shutdown 192.168.1.10`	Shuts down the peer. By default, the peer is enabled when it is defined.	
ip msdp password *peer-ip-address password* **Example:** `switch(config)# ip msdp password 192.168.1.10 my_md5_password`	Enables an MD5 password for the peer. By default, no MD5 password is enabled.	
ip msdp sa-policy *peer-ip-address policy-name* **in** **Example:** `switch(config)# ip msdp sa-policy 192.168.1.10 my_incoming_sa_policy in`	Enables a route-map policy for incoming SA messages. By default, all SA messages are received.	
ip msdp sa-policy *peer-ip-address policy-name* **out** **Example:** `switch(config)# ip msdp sa-policy 192.168.1.10 my_outgoing_sa_policy out`	Enables a route-map policy for outgoing SA messages. By default, all registered sources are sent in SA messages.	
ip msdp sa-limit *peer-ip-address limit* **Example:** `switch(config)# ip msdp sa-limit 192.168.1.10 5000`	Sets a limit on the number of (S, G) entries accepted from the peer. By default, there is no limit.	

	Command or Action	Purpose
Step 3	(Optional) **show ip msdp peer** [*peer-address*] [**vrf** [*vrf-name* | **all**]] **Example:** `switch(config)# show ip msdp peer 192.168.1.10`	Displays detailed MDSP peer information.
Step 4	(Optional) **copy running-config startup-config** **Example:** `switch(config)# copy running-config startup-config`	Copies the running configuration to the startup configuration.

Configuring MSDP Global Parameters

You can configure the optional MSDP global parameters described in this table.

Table 19: MSDP Global Parameters

Parameter	Description
Originator interface name	IP address used in the RP field of an SA message entry. When Anycast RPs are used, all RPs use the same IP address. You can use this parameter to define a unique IP address for the RP of each MSDP peer. By default, the software uses the RP address of the local system. **Note** We recommend that you use a loopback interface for the RP address.
Group limit	Maximum number of (S, G) entries that the software creates for the specified prefix. The software ignores groups when the group limit is exceeded and logs a violation. By default, no group limit is defined.
SA interval	Interval at which the software transmits Source-Active (SA) messages. The range is from 60 to 65,535 seconds. The default is 60 seconds.

Before you begin

Ensure that you have installed the Enterprise Services license and enabled PIM and MSDP.

Procedure

<table>
<tr><th></th><th>Command or Action</th><th>Purpose</th></tr>
<tr><td>Step 1</td><td>configure terminal
Example:
<pre>switch# configure terminal
switch(config)#</pre></td><td>Enters global configuration mode.</td></tr>
<tr><td>Step 2</td><td>
<table>
<tr><th>Option</th><th>Description</th></tr>
<tr><td>ip msdp originator-id interface
Example:
<pre>switch(config)# ip msdp originator-id loopback0</pre></td><td>Sets a description string for the peer. By default, the peer has no description.
Sets the IP address used in the RP field of an SA message entry. By default, the software uses the RP address of the local system.
Note We recommend that you use a loopback interface for the RP address.</td></tr>
<tr><td>ip msdp group-limit limit source source-prefix
Example:
<pre>switch(config)# ip msdp group-limit 1000 source 192.168.1.0/24</pre></td><td>Maximum number of (S, G) entries that the software creates for the specified prefix. The software ignores groups when the group limit is exceeded and logs a violation. By default, no group limit is defined.</td></tr>
<tr><td>ip msdp sa-interval seconds
Example:
<pre>switch(config)# ip msdp sa-interval 80</pre></td><td>Interval at which the software transmits Source-Active (SA) messages. The range is from 60 to 65,535 seconds. The default is 60 seconds.</td></tr>
</table>
</td><td></td></tr>
</table>

	Command or Action	Purpose
Step 3	(Optional) **show ip msdp summary** [**vrf** [*vrf-name* | **all**]] **Example:** `switch(config)# show ip msdp summary`	Displays a summary of the MDSP configuration.
Step 4	(Optional) **copy running-config startup-config** **Example:** `switch(config)# copy running-config startup-config`	Copies the running configuration to the startup configuration.

Configuring MSDP Mesh Groups

You can configure optional MDSP mesh groups in global configuration mode by specifying each peer in the mesh. You can configure multiple mesh groups on the same router and multiple peers per mesh group.

Before you begin

Ensure that you have installed the Enterprise Services license and enabled PIM and MSDP.

Procedure

	Command or Action	Purpose

Step 1	**configure terminal**
	Example:
	`switch# configure terminal`
	`switch(config)#`
		Enters global configuration mode.

Step 2	**ip msdp mesh-group** *peer-ip-addr mesh-name* **Example:** `switch(config)# ip msdp mesh-group 192.168.1.10 my_mesh_1`	Configures an MSDP mesh with the peer IP address specified. You can configure multiple meshes on the same router and multiple peers per mesh group. By default, no mesh groups are configured.
Step 3	Repeat Step 2 for each MSDP peer in the mesh by changing the peer IP address.	—
Step 4	(Optional) **show ip msdp mesh-group** [*mesh-group*] [**vrf** [*vrf-name* | **all**]] **Example:** `switch# show ip msdp mesh-group`	Displays information about the MDSP mesh group configuration.
Step 5	(Optional) **copy running-config startup-config** **Example:** `switch(config)# copy running-config startup-config`	Copies the running configuration to the startup configuration.

Restarting the MSDP Process

Before you begin

You can restart the MSDP process and optionally flush all routes.

Procedure

	Command or Action	Purpose
Step 1	**restart msdp** **Example:** `switch# restart msdp`	Restarts the MSDP process.

Step 2	**configure terminal**
	Example:
	`switch# configure terminal`
	`switch(config)#`
		Enters global configuration mode.

Step 3	**ip msdp flush-routes** **Example:** `switch(config)# ip msdp flush-routes`	Removes routes when the MSDP process is restarted. By default, routes are not flushed.
Step 4	(Optional) **show running-configuration | include flush-routes** **Example:** `switch(config)# show running-configuration | include flush-routes`	Displays flush-routes configuration lines in the running configuration.
Step 5	(Optional) **copy running-config startup-config** **Example:** `switch(config)# copy running-config startup-config`	Copies the running configuration to the startup configuration.

Verifying the MSDP Configuration

To display the MSDP configuration information, perform one of the following tasks.

Command	Description
show ip msdp count [*as-number*] [**vrf** [*vrf-name* | **all**]]	Displays MSDP (S, G) entry and group counts by the autonomous system (AS) number.
show ip msdp mesh-group [*mesh-group*] [**vrf** [*vrf-name* | **all**]]	Displays the MSDP mesh group configuration.

Command	Description
show ip msdp peer [*peer-address*] [**vrf** [*vrf-name* | **all**]]	Displays MSDP information for the MSDP peer.
show ip msdp rpf [*rp-address*] [**vrf** [*vrf-name* | **all**]]	Displays the next-hop AS on the BGP path to an RP address.
show ip msdp sources [**vrf** [*vrf-name* | **all**]]	Displays the MSDP-learned sources and violations of configured group limits.
show ip msdp summary [**vrf** [*vrf-name* | **all**]]	Displays a summary of the MSDP peer configuration.

Monitoring MSDP

You can display and clear MSDP statistics by using the features in this section.

Displaying Statistics

You can display MSDP statistics using these commands.

Command	Description
show ip msdp policy statistics sa-policy *peer-address* {**in** | **out**} [**vrf** [*vrf-name* | **all**]]	Displays the MSDP policy statistics for the MSDP peer.
show ip msdp {**sa-cache** | **route**} [*source-address*] [*group-address*] [**vrf** [*vrf-name* | **all**]] [*asn-number*] [**peer** *peer-address*]	Displays the MSDP SA route cache. If you specify the source address, all groups for that source are displayed. If you specify a group address, all sources for that group are displayed.

Clearing Statistics

You can clear the MSDP statistics using these commands.

Command	Description
clear ip msdp peer [*peer-address*] [**vrf** *vrf-name*]	Clears the TCP connection to an MSDP peer.
clear ip msdp policy statistics sa-policy *peer-address* {**in** | **out**} [**vrf** *vrf-name*]	Clears statistics counters for MSDP peer SA policies.
clear ip msdp statistics [*peer-address*] [**vrf** *vrf-name*]	Clears statistics for MSDP peers.
clear ip msdp {**sa-cache** | **route**} [*group-address*] [**vrf** [*vrf-name* | **all**]]	Clears the group entries in the SA cache.

Configuration Examples for MSDP

To configure MSDP peers, some of the optional parameters, and a mesh group, follow these steps for each MSDP peer:

1. Configure the MSDP peering relationship with other routers.

```
switch# configure terminal
switch(config)# ip msdp peer 192.168.1.10 connect-source ethernet 1/0 remote-as 8
```

2. Configure the optional peer parameters.

```
switch# configure terminal
switch(config)# ip msdp password 192.168.1.10 my_peer_password_AB
```

3. Configure the optional global parameters.

```
switch# configure terminal
switch(config)# ip msdp sa-interval 80
```

4. Configure the peers in each mesh group.

```
switch# configure terminal
switch(config)# ip msdp mesh-group 192.168.1.10 mesh_group_1
```

The following example shows how to configure a subset of the MSDP peering that is shown below.

RP 3: 192.168.3.10 (AS 7)

```
configure terminal
  ip msdp peer 192.168.1.10 connect-source ethernet 1/1
  ip msdp peer 192.168.2.10 connect-source ethernet 1/2
  ip msdp peer 192.168.6.10 connect-source ethernet 1/3 remote-as
9
  ip msdp password 192.168.6.10 my_peer_password_36
  ip msdp sa-interval 80
  ip msdp mesh-group 192.168.1.10 mesh_group_123
  ip msdp mesh-group 192.168.2.10 mesh_group_123
  ip msdp mesh-group 192.168.3.10 mesh_group_123
```

RP 5: 192.168.5.10 (AS 8)

```
configure terminal
  ip msdp peer 192.168.4.10 connect-source ethernet 1/1
  ip msdp peer 192.168.6.10 connect-source ethernet 1/2 remote-as
9
  ip msdp password 192.168.6.10 my_peer_password_56
  ip msdp sa-interval 80
```

RP 6: 192.168.6.10 (AS 9)

```
configure terminal
  ip msdp peer 192.168.7.10 connect-source ethernet 1/1
  ip msdp peer 192.168.3.10 connect-source ethernet 1/2 remote-as
7
  ip msdp peer 192.168.5.10 connect-source ethernet 1/3 remote-as
8
  ip msdp password 192.168.3.10 my_peer_password_36
  ip msdp password 192.168.5.10 my_peer_password_56
  ip msdp sa-interval 80
```

Related Documents

Related Topic	Document Title
Configuring MBGP	*CN93240YC-FX2 NX-OS Unicast Routing Configuration Guide*

Standards

Standards	Title
RFC 4624	Multicast Source Discovery Protocol (MSDP) MIB

CHAPTER 9

Configuring MVR

This chapter describes how to configure the MVR feature on NX-OS devices. This

chapter contains the following sections:

About MVR

In a typical Layer 2 multi-VLAN network, subscribers to a multicast group can be on multiple VLANs. To maintain data isolation between these VLANs, the multicast stream on the source VLAN must be passed to a router, which replicates the stream on all subscriber VLANs, wasting upstream bandwidth.

Multicast VLAN registration (MVR) allows a Layer 2 switch to forward the multicast data from a source on a common assigned VLAN to the subscriber VLANs, conserving upstream bandwidth by bypassing the router. The switch forwards multicast data for MVR IP multicast streams only to MVR ports on which hosts have joined, either by IGMP reports or by MVR static configuration. The switch forwards IGMP reports received from MVR hosts only to the source port. For other traffic, VLAN isolation is preserved.

MVR requires at least one VLAN to be designated as the common VLAN to carry the multicast stream from the source. More than one such multicast VLAN (MVR VLAN) can be configured in the system, and you can configure a global default MVR VLAN as well as interface-specific default MVR VLANs. Each multicast group using MVR is assigned to an MVR VLAN.

MVR allows a subscriber on a port to subscribe and unsubscribe to a multicast stream on the MVR VLAN by sending IGMP join and leave messages. IGMP leave messages from an MVR group are handled according to the IGMP configuration of the VLAN on which the leave message is received. If IGMP fast leave is enabled on the VLAN, the port is removed immediately; otherwise, an IGMP query is sent to the group to determine whether other hosts are present on the port.

MVR Interoperation with Other Features

MVR and IGMP Snooping

Although MVR operates on the underlying mechanism of IGMP snooping, the two features operate independently of each other. One feature can be enabled or disabled without affecting the operation of the other feature. If IGMP snooping is disabled globally or on a VLAN and MVR is enabled on the VLAN, IGMP snooping is internally enabled on the VLAN. Joins received for MVR groups on non-MVR receiver ports or joins received for non-MVR groups on MVR receiver ports are processed by IGMP snooping.

MVR and vPCs

- As with IGMP snooping, IGMP control messages received by virtual port channel (vPC) peer switches are exchanged between the peers, allowing synchronization of MVR group information.
- MVR configuration must be consistent between the peers.
- The **no ip igmp snooping mrouter vpc-peer-link** command applies to MVR. With this command, multicast traffic is not sent to a peer link for the source VLAN and receiver VLAN unless an orphan port is in the VLAN.
- The **show mvr member** command shows the multicast group on the vPC peer switch. However, the vPC peer switch does not show the multicast groups if it does not receive the IGMP membership report of the groups.

Licensing Requirements for MVR

The following table shows the licensing requirements for this feature:

Product	License Requirement
NX-OS	This feature does not require a license. Any feature not included in a license package is bundled with the nx-os image and is provided at no extra charge to you. For a complete explanation of the NX-OS licensing scheme, see the *NX-OS Licensing Guide*.

Guidelines and Limitations for MVR

MVR has the following guidelines and limitations:

- MVR is supported only on Layer 2 Ethernet ports, such as individual ports, port channels, and virtual Ethernet (vEth) ports.
- MVR receiver ports can only be access ports; they cannot be trunk ports. MVR source ports can be either access or trunk ports.=

- MVR configuration on Flex Link ports is not supported.
- Priority tagging is not supported on MVR receiver ports.
- The total number of MVR VLANs cannot exceed 250.

Default MVR Settings

This table lists the default settings for MVR parameters.

Table 20: Default MVR Parameters

Parameter	Default
MVR	Disabled globally and per interface
Global MVR VLAN	None configured
Interface (per port)	Neither a receiver nor a source port

Configuring MVR

Configuring MVR Global Parameters

You can globally enable MVR and various configuration parameters.

Procedure

	Command or Action	Purpose

Step 1	**configure terminal**
	Example:
	`switch# configure terminal`
	`switch(config)#`
		Enters global configuration mode.

Step 2	[**no**]**mvr**
	Example:
	`switch(config)# mvr`
	`switch(config-mvr)#`
		Globally enables MVR. The default is disabled.
		Use the **no** form of the command to disable MVR.

Step 3	[**no**] **mvr-vlan** *vlan-id* **Example:** `switch(config-mvr)# mvr-vlan 7`	Specifies the global default MVR VLAN. The MVR VLAN is the source of the multicast message that subsequent receivers subscribe to. The range is from 1 to 4094. Use the **no** form of the command to clear the MVR VLAN.

	Command or Action	Purpose
Step 4	[**no**] **mvr-group** *addr* [/*mask*] [**count** *groups*] [**vlan** *vlan-id*] **Example:** `switch(config-mvr)# mvr-group 230.1.1.1 count 4`	Adds a multicast group at the specified IPv4 address (and optional netmask length) to the global default MVR VLAN. You can repeat this command to add additional groups to the MVR VLAN. The IP address is entered in the format *a.b.c.d/m*, where *m* is the number of bits in the netmask, from 1 to 31. You can optionally specify a number of MVR groups using contiguous multicast IP addresses starting with the specified IP address. Use the **count** keyword followed by a number from 1 to 64. You can optionally specify an MVR VLAN for the group by using the **vlan** keyword. Otherwise, the group is assigned to the default MVR VLAN. Use the **no** form of the command to clear the group configuration.
Step 5	(Optional) **clear mvr counters** [**source-ports** | **receiver-ports**] **Example:** `switch(config-mvr)# clear mvr counters`	Clears MVR IGMP packet counters.
Step 6	(Optional) **show mvr** **Example:** `switch(config-mvr)# show mvr`	Displays the global MVR configuration.
Step 7	(Optional) **copy running-config startup-config** **Example:** `switch(config-mvr)# copy running-config startup-config`	Copies the running configuration to the startup configuration.

Configuring MVR Interfaces

You can configure MVR interfaces on your NX-OS device.

Procedure

	Command or Action	Purpose
Step 1	**configure terminal** **Example:**	Enters global configuration mode.

	Command or Action	Purpose
	`switch# configure terminal` `switch(config)#`	

Step 2	**mvr**
	Example:
	`switch(config)# mvr`
	`switch(config-mvr)#`
		Globally enables MVR. The default is disabled.
		Note If MVR is enabled globally, this command is not required.

Step 3	**interface** {**ethernet** *slot/port* | **port-channel** *channel-number* | **vethernet** *number*}
	Example:
	`switch(config-mvr)# interface ethernet 2/2`
	`switch(config-mvr-if)#`
		Specifies the Layer 2 port to configure and enters interface configuration mode.

Step 4	[**no**] **mvr-type** {**source** | **receiver**} **Example:** `switch(config-mvr-if)# mvr-type source`	Configures an MVR port as one of these types of ports: • **source**—An uplink port that sends and receives multicast data is configured as an MVR source. The port automatically becomes a static receiver of MVR multicast groups. A source port should be a member of the MVR VLAN. • **receiver**—An access port that is connected to a host that wants to subscribe to an MVR multicast group is configured as an MVR receiver. A receiver port receives data only when it becomes a member of the multicast group by using IGMP leave and join messages. If you attempt to configure a non-MVR port with MVR characteristics, the configuration is cached and does not take effect until the port becomes an MVR port. The default port mode is non-MVR.
Step 5	(Optional) [**no**] **mvr-vlan** *vlan-id* **Example:** `switch(config-mvr-if)# mvr-vlan 7`	Specifies an interface default MVR VLAN that overrides the global default MVR VLAN for joins received on the interface. The MVR VLAN is the source of the multicast message that subsequent receivers subscribe to. The range is from 1 to 4094.
Step 6	(Optional) [**no**] **mvr-group** *addr* [/*mask*] [**vlan** *vlan-id*] **Example:** `switch(config-mvr-if)# mvr-group 225.1.3.1 vlan 100`	Adds a multicast group at the specified IPv4 address (and optional netmask length) to the interface MVR VLAN, overriding the global MVR group configuration. You can repeat this command to add additional groups to the MVR.

	Command or Action	Purpose
		The IP address is entered in the format *a.b.c.d/m*, where *m* is the number of bits in the netmask, from 1 to 31. You can optionally specify an MVR VLAN for the group by using the **vlan** keyword; otherwise, the group is assigned to the interface default (if specified) or the global default MVR VLAN. Use the **no** form of the command to clear the IPv4 address and netmask.
Step 7	(Optional) **copy running-config startup-config** **Example:** `switch(config-mvr-if)# copy running-config startup-config`	Copies the running configuration to the startup configuration.

Suppressing IGMP Query Forwarding from VLANs

To suppress the IGMP general query from the source VLAN to the receiver VLAN perform the following steps.

Procedure

	Command or Action	Purpose

Step 1	**configure terminal**
	Example:
	`switch# configure terminal`
	`switch(config)#`
		Enters global configuration mode.

Step 2	**mvr-config**
	Example:
	`switch# mvr-config`
	`switch(config-mvr)#`
		Enters global MVR configuration mode.

Step 3	**mvr-suppress-query vlan** *vlan-ID*
	Example:
	`switch(config-mvr)# mvr-suppress-query vlan 1-5`
	`switch(config-mvr)#`
		Displays the MVR ID or source VLAN range from where the general queries need to be suppressed. The VLAN ID value is 1 to 3967. The VLAN ID may also be expressed as a range 1-5, 10 or 2-5, 7-19.

Verifying the MVR Configuration

To display the MVR configuration information, perform one of the following tasks:

Command	Description
show mvr	Displays the MVR subsystem configuration and status.
show mvr groups	Displays the MVR group configuration.
show ip igmp snooping [**vlan** *vlan-id*]	Displays information about IGMP snooping on the specified VLAN.
show mvr interface {**ethernet** *slot/port* | **port-channel** *number*}	Displays the MVR configuration on the specified interface.
show mvr members [**count**]	Displays the number and details of all MVR receiver members.
show mvr members interface {**ethernet** *slot/port* | **port-channel** *number*}	Displays details of MVR members on the specified interface.
show mvr members vlan *vlan-id*	Displays details of MVR members on the specified VLAN.
show mvr receiver-ports [**ethernet** *slot/port* | **port-channel** *number*]	Displays all MVR receiver ports on all interfaces or on the specified interface.
show mvr source-ports [**ethernet** *slot/port* | **port-channel** *number*]	Displays all MVR source ports on all interfaces or on the specified interface.

This example shows how to verify the MVR parameters:

```
switch# show mvr
MVR Status          : enabled
Global MVR VLAN     : 100
Number of MVR VLANs : 4
```

This example shows how to verify the MVR group configuration:

```
switch# show mvr groups
* - Global default MVR VLAN.

Group start    Group end       Count  MVR-VLAN Interface
                               Mask
-------------  --------------- ------ -------- -----------
228.1.2.240    228.1.2.255     /28     101
230.1.1.1      230.1.1.4        4     *100
235.1.1.6      235.1.1.6        1      340
225.1.3.1      225.1.3.1        1     *100     Eth1/10
```

This example shows how to verify the MVR interface configuration and status:

```
switch# show mvr interface
Port         VLAN Type      Status    MVR-VLAN
----         ---- ----      ------    --------
Po10         100  SOURCE    ACTIVE    100-101
Po201        201  RECEIVER  ACTIVE    100-101,340
Po202        202  RECEIVER  ACTIVE    100-101,340
Po203        203  RECEIVER  ACTIVE    100-101,340
Po204        204  RECEIVER  INACTIVE  100-101,340
Po205        205  RECEIVER  ACTIVE    100-101,340
Po206        206  RECEIVER  ACTIVE    100-101,340
```

```
Po207         207  RECEIVER  ACTIVE    100-101,340
Po208         208  RECEIVER  ACTIVE    2000-2001
Eth1/9        340  SOURCE    ACTIVE    340
Eth1/10       20   RECEIVER  ACTIVE    100-101,340
Eth2/2        20   RECEIVER  ACTIVE    100-101,340
Eth102/1/1    102  RECEIVER  ACTIVE    100-101,340
Eth102/1/2    102  RECEIVER  INACTIVE  100-101,340
Eth103/1/1    103  RECEIVER  ACTIVE    100-101,340
Eth103/1/2    103  RECEIVER  ACTIVE    100-101,340

Status INVALID indicates one of the following misconfiguration:
a) Interface is not a switchport.
b) MVR receiver is not in access mode.
c) MVR source is in fex-fabric mode.
```

This example shows how to display all MVR members:

```
switch# show mvr members
MVR-VLAN  Group Address  Status   Members
--------  -------------  -------  -------
100       230.1.1.1      ACTIVE    Po201 Po202 Po203 Po205 Po206
100       230.1.1.2      ACTIVE    Po205 Po206 Po207 Po208
340       235.1.1.6      ACTIVE    Eth102/1/1
101       225.1.3.1      ACTIVE    Eth1/10 Eth2/2
101       228.1.2.241    ACTIVE    Eth103/1/1 Eth103/1/2
```

This example shows how to display all MVR receiver ports on all interfaces:

```
switch# show mvr receiver-ports
Port          MVR-VLAN  Status   Joins         Leaves
                                 (v1,v2,v3)
------------  --------  -------- ------------  ------------
Po201         100       ACTIVE   8             2
Po202         100       ACTIVE   8             2
Po203         100       ACTIVE   8             2
Po204         100       INACTIVE 0             0
Po205         100       ACTIVE   10            6
Po206         100       ACTIVE   10            6
Po207         100       ACTIVE   5             0
Po208         100       ACTIVE   6             0
Eth1/10       101       ACTIVE   12            2
Eth2/2        101       ACTIVE   12            2
Eth102/1/1    340       ACTIVE   16            15
Eth102/1/2    340       INACTIVE 16            16
Eth103/1/1    101       ACTIVE   33            0
Eth103/1/2    101       ACTIVE   33            0
```

This example shows how to display all MVR source ports on all interfaces:

```
switch# show mvr source-ports
Port          MVR-VLAN  Status
------------  --------  --------
Po10          100       ACTIVE
Eth1/9        340       ACTIVE
```

Configuration Examples for MVR

The following example shows how to globally enable MVR and configure the global parameters:

```
switch# configure terminal
switch(config)# mvr
switch(config-mvr)# mvr-vlan 100
switch(config-mvr)# mvr-group 230.1.1.1 count 4
switch(config-mvr)# mvr-group 228.1.2.240/28 vlan 101
switch(config-mvr)# mvr-group 235.1.1.6 vlan 340

switch# show mvr
MVR Status           : enabled
Global MVR VLAN      : 100
Number of MVR VLANs  : 3
```

The following example shows how to configure an Ethernet port as an MVR receiver port:

```
switch# configure terminal
switch(config)# mvr
switch(config-mvr)# interface ethernet 1/10
switch(config-mvr-if)# mvr-group 225.1.3.1 vlan 100
switch(config-mvr-if)# mvr-type receiver
switch(config-mvr-if)## copy running-config startup-config
```

CHAPTER 10

Configuring Microsoft Network Load Balancing (NLB)

This chapter describes how to configure the Microsoft Network Load Balancing (NLB) feature on Cisco NX-OS devices.

About Network Load Balancing (NLB)

Network Load Balancing (NLB) technology is used to distribute client requests across a set of servers. There are three primary modes of NLB: unicast, multicast, and Internet Group Management Protocol (IGMP) multicast:

- **Unicast mode** assigns the cluster a virtual IP and virtual MAC address. This method relies on unknown unicast flooding. Because the virtual MAC address is not learned on any switchports, traffic that is destined to the virtual MAC address is flooded within the VLAN. This means that all clustered servers receive traffic destined to the virtual MAC address. One downside to this method is that all devices in the VLAN receive this traffic. The only way to mitigate this behavior is to limit the NLB VLAN to only the NLB server interfaces in order to avoid flooding to interfaces that should receive the traffic.
- **Multicast mode** assigns a unicast IP address to a non-Internet Assigned Numbers Authority (IANA) multicast MAC address (03xx.xxxx.xxxx). IGMP snooping does not dynamically program this address, which results in flooding of the NLB traffic in the VLAN. Not reqiring a PIM-enabled SVI or the IGMP snooping querier means that NLB works with custom non-IP multicast applications. For more information see, Multicast Mode, on page 173
- **IGMP multicast mode** assigns the cluster a virtual unicast IP address and a virtual multicast MAC address within the IANA range (01:00:5E:XX:XX:XX). The clustered servers send IGMP joins for the configured multicast group, and thus the switch dynamically populates its IGMP snooping table to point toward the clustered servers, which prevents unicast flooding. See IGMP Multicast Mode, on page 174 for configuration examples.

This section describes how to configure a CN93240YC-FX2 switches for multicast and IGMP multicast mode NLB. As previously referenced, multicast NLB requires that you have a unicast IP address that is mapped to a multicast MAC address.

- Static Address Resolution Protocol (ARP) multicast.
- MAC address to a unicast IP address, but the traffic to that IP address floods the VLAN.

Licensing Requirements for NLB

Product	License Requirement
NX-OS	NLB requires no license. Any feature not included in a license package is bundled with the nx-os image and is provided at no extra charge to you. For a complete explanation of the NX-OS licensing scheme, see the NX-OS Licensing Guide.

Guidelines and Limitations for NLB

Network Load Balancing (NLB) has the following configuration guidelines and limitations:

- NLB is supported on CN93240YC-FX2 platform switches.
- FEX HIF interfaces cannot receive a multicast NLB flow.
- If none of the ports in the interface set is UP, the traffic floods to all ports in the VLAN.
- L2 and L3 regular multicast is not supported to, from or inside the NLB VLAN.
- NLB traffic that enters the NLB VLAN may be looped back to the source interface. This looped back NLB traffic time-to-live (TTL) is decremented even though it is intra-VLAN.
- Multicast Mode - If servers/firewalls move, the administrator must update the static multicast MAC table configuration.
- IGMP Multicast Mode - If servers/firewalls move, the administrator must update the static-group configuration.=

Prerequisites for Microsoft Network Load Balancing (NLB)

Microsoft Network Load Balancing (NLB) has the following prerequisites:

- You are logged into the device.
- For global configuration commands, you are in the correct virtual routing and forwarding (VRF) mode. The default configuration mode shown in the examples in this chapter applies to the default VRF.
- Multicast NLB requires that you have a unicast IP address mapped to a multicast MAC address.

Multicast Mode

Multicast mode assigns a unicast IP address to a non-Internet Assigned Numbers Authority (IANA) multicast MAC address (03xx.xxxx.xxxx). IGMP snooping does not dynamically program this address, which results in flooding of the NLB traffic in the VLAN. Refer to Option 2A for an example of how to configure for this mode. The following example shows how to configure for IGMP Multicast Mode:

Example 1: Static ARP + MAC-based L2 Multicast Lookups + Static Joins + Non-IP Multicast MAC

This option does not require a PIM-enabled SVI or the IGMP snooping querier; works with non-IP multicast applications (custom applications).

Note The **hardware profile multicast nlb** CLI must be enabled on the switch to support Multicast Mode.

1. Configure a static ARP entry that maps the unicast IP address to a multicast MAC address, but this time in the non-IP address multicast range:

```
interface Vlan10
no shutdown
ip address 10.1.2.1/24
ip arp 10.1.2.200 03bf.0000.1111
```

2. Enable MAC-based Layer 2 multicast lookups in the VLAN (by default, multicast lookups are based on the destination multicast IP address):

Note You must use MAC-based lookups in VLANs where you want to constrain IP address unicast packets with multicast MAC addresses.

```
vlan configuration 10
layer-2 multicast lookup mac
```

3. Configure static MAC address-table entries that point to the interfaces connected to the NLB server and any redundant interface:

```
mac address-table multicast 03bf.0000.1111 vlan 10 interface Ethernet8/2
mac address-table multicast 03bf.0000.1111 vlan 10 interface Ethernet8/4
mac address-table multicast 03bf.0000.1111 vlan 10 interface Ethernet8/7
```

IGMP Multicast Mode

IGMP multicast mode assigns the cluster a virtual unicast IP address and a virtual multicast MAC address within the IANA range (01:00:5E:XX:XX:XX). The clustered servers send IGMP joins for the configured multicast group, and thus the switch dynamically populates its IGMP snooping table to point toward the clustered servers, which prevents unicast flooding. The following describes three examples of how to configure for IGMP Multicast Mode:

Option 1: Static ARP + MAC-based L2 Multicast Lookups + Dynamic Joins

This option allows servers and firewalls to dynamically join or leave the corresponding group; enables or disables reception of the target traffic (for example, maintenance mode).

Note The **hardware profile multicast nlb** CLI must be enabled on the switch to support IGMP Multicast Mode.

1. Configure a static ARP entry that maps the unicast IP address to a multicast MAC address in the IP address multicast range on a Protocol Independent Multicast (PIM)-enabled interface:

```
interface Vlan10
no shutdown
ip address 10.1.2.1/24
ip pim sparse-mode
ip arp 10.1.2.200 0100.5E01.0101
```

2. Enable MAC-based Layer 2 multicast lookups in the VLAN (by default, multicast lookups are based on the destination multicast IP address):

```
vlan configuration 10
layer-2 multicast lookup mac
```

Option 2: Static ARP + MAC-based L2 Multicast Lookups + Dynamic Joins with IGMP Snooping Querier

Option 2 does not require PIM-enabled SVI and allows servers and firewalls to dynamically join or leave the corresponding group; enables or disables reception of the target traffic (for example, maintenance mode).

Note The **hardware profile multicast nlb** CLI must be enabled on the switch to support IGMP Multicast Mode.

1. Configure a static ARP entry like in Option 1, but do not enable PIM on the switch virtual interface (SVI).

```
interface Vlan10
no shutdown
ip address 10.1.2.1/24
ip arp 10.1.2.200 0100.5E01.0101
```

2. Enable MAC-based Layer 2 multicast lookups in the VLAN, and enable the Internet Group Management Protocol (IGMP) snooping querier:

```
vlan configuration 10
ip igmp snooping querier 10.1.1.254
layer-2 multicast lookup mac
```

Option 3: Static ARP + MAC-based L2 Multicast Lookups + Static Joins + IP Multicast MAC

Option three does not require a PIM-enabled SVI or the IGMP snooping querier.

Note The **hardware profile multicast nlb** CLI must be enabled on the switch to support IGMP Multicast Mode.

1. Configure a static ARP entry that maps the unicast IP address to a multicast MAC address in the IP address multicast range:

```
interface Vlan10
no shutdown
ip address 10.1.2.1/24
ip arp 10.1.2.200 0100.5E01.0101
```

2: Enable MAC-based Layer 2 multicast lookups in the VLAN (by default, multicast lookups are based on the destination multicast IP address):

```
vlan configuration 10
layer-2 multicast lookup mac
```

You must use MAC-based lookups in VLANs where you want to constrain IP address unicast packets with multicast MAC addresses.

3. Configure static IGMP snooping group entries for the interfaces connected to the NLB server that needs the traffic:

```
vlan configuration 10
ip igmp snooping static-group 239.1.1.1 interface Ethernet8/2
ip igmp snooping static-group 239.1.1.1 interface Ethernet8/4
ip igmp snooping static-group 239.1.1.1 interface Ethernet8/7
```

Verifying the NLB Configuration

To display the NLB configuration information, perform one of the following tasks.

Command	Description
show ip arp *virtual-address*	Displays the ARP table.
show ip igmp snooping groups [*source* [*group*] | *group* [*source*]] [**vlan** *vlan-id*] [**detail**]	Displays IGMP snooping information about groups by VLAN.
show ip igmp snooping mac-oif vlan *vlan-id*	Displays IGMP snooping static MAC addresses.

APPENDIX A

IETF RFCs for IP Multicast

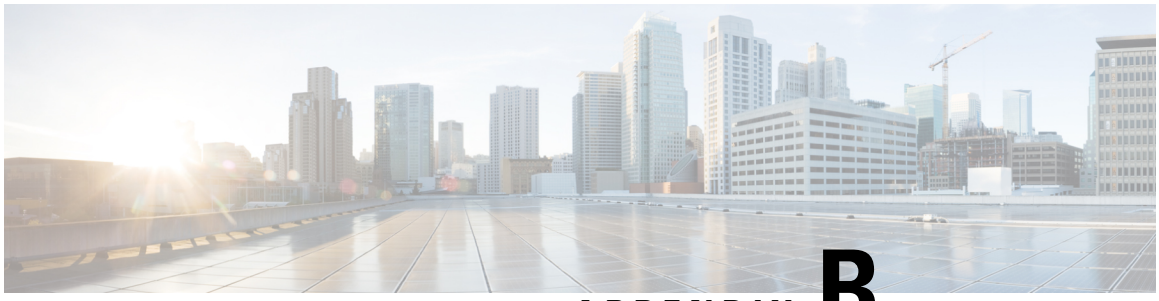
This appendix contains Internet Engineering Task Force (IETF) RFCs related to IP multicast. For information about IETF RFCs, see https://www.ietf.org/search/?query=RFC.

IETF RFCs for IP Multicast

This table lists the RFCs related to IP multicast.

RFCs	Title
RFC 2236	*Internet Group Management Protocol*
RFC 2365	*Administratively Scoped IP Multicast*
RFC 2710	*Multicast Listener Discovery (MLD) for IPv6*
RFC 2858	*Multiprotocol Extensions for BGP-4*
RFC 3376	*Internet Group Management Protocol*
RFC 3446	*Anycast Rendezvous Point (RP) mechanism using Protocol Independent Multicast (PIM) and Multicast Source Discovery Protocol (MSDP)*
RFC 3569	*An Overview of Source-Specific Multicast (SSM)*
RFC 3618	*Multicast Source Discovery Protocol (MSDP)*
RFC 3810	*Multicast Listener Discovery Version 2 (MLDv2) for IPv6*
RFC 4601	*Protocol Independent Multicast - Sparse Mode (PIM-SM): Protocol Specification (Revised)*
RFC 4610	*Anycast-RP Using Protocol Independent Multicast (PIM)*
RFC 5132	*IP Multicast MIB*

IETF RFCs for IP Multicast



APPENDIX B

Configuration Limits for NX-OS Multicast

This appendix describes the configuration limits for NX-OS multicast.

Configuration Limits

The features supported by NX-OS have maximum configuration limits. Some of the features have configurations that support limits less than the maximum limits.

The configuration limits are documented in the CN93240YC-FX2 NX-OS Verified Scalability Guide.

Configuration Limits for Cisco NX-OS Multicast