Cisco Nexus 9000 Series NX-OS Multicast Routing Configuration Guide, Release 10.1(x)

First Published: 2021-02-16

Last Modified: 2023-09-11

Americas Headquarters
Cisco Systems, Inc.
170 West Tasman Drive
San Jose, CA 95134-1706
USA
http://www.cisco.com
Tel: 408 526-4000
800 553-NETS (6387)
Fax: 408 527-0883

THE SPECIFICATIONS AND INFORMATION REGARDING THE PRODUCTS REFERENCED IN THIS DOCUMENTATION ARE SUBJECT TO CHANGE WITHOUT NOTICE. EXCEPT AS MAY OTHERWISE BE AGREED BY CISCO IN WRITING, ALL STATEMENTS, INFORMATION, AND RECOMMENDATIONS IN THIS DOCUMENTATION ARE PRESENTED WITHOUT WARRANTY OF ANY KIND, EXPRESS OR IMPLIED.

The Cisco End User License Agreement and any supplemental license terms govern your use of any Cisco software, including this product documentation, and are located at: https://www.cisco.com/c/en/us/about/legal/cloud-and-software/software-terms.html. Cisco product warranty information is available at https://www.cisco.com/c/en/us/products/warranty-listing.html. US Federal Communications Commission Notices are found here https://www.cisco.com/c/en/us/products/us-fcc-notice.html.

IN NO EVENT SHALL CISCO OR ITS SUPPLIERS BE LIABLE FOR ANY INDIRECT, SPECIAL, CONSEQUENTIAL, OR INCIDENTAL DAMAGES, INCLUDING, WITHOUT LIMITATION, LOST PROFITS OR LOSS OR DAMAGE TO DATA ARISING OUT OF THE USE OR INABILITY TO USE THIS MANUAL, EVEN IF CISCO OR ITS SUPPLIERS HAVE BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES.

Any products and features described herein as in development or available at a future date remain in varying stages of development and will be offered on a when-and if-available basis. Any such product or feature roadmaps are subject to change at the sole discretion of Cisco and Cisco will have no liability for delay in the delivery or failure to deliver any products or feature roadmap items that may be set forth in this document.

Any Internet Protocol (IP) addresses and phone numbers used in this document are not intended to be actual addresses and phone numbers. Any examples, command display output, network topology diagrams, and other figures included in the document are shown for illustrative purposes only. Any use of actual IP addresses or phone numbers in illustrative content is unintentional and coincidental.

The documentation set for this product strives to use bias-free language. For the purposes of this documentation set, bias-free is defined as language that does not imply discrimination based on age, disability, gender, racial identity, ethnic identity, sexual orientation, socioeconomic status, and intersectionality. Exceptions may be present in the documentation due to language that is hardcoded in the user interfaces of the product software, language used based on RFP documentation, or language that is used by a referenced third-party product.

Cisco and the Cisco logo are trademarks or registered trademarks of Cisco and/or its affiliates in the U.S. and other countries. To view a list of Cisco trademarks, go to this URL: https://www.cisco.com/c/en/us/about/legal/trademarks.html. Third-party trademarks mentioned are the property of their respective owners. The use of the word partner does not imply a partnership relationship between Cisco and any other company. (1721R)

© 2021 Cisco Systems, Inc. All rights reserved.

CONTENTS

PREFACE

Preface xi

Audience xi

Document Conventions xi

Related Documentation for Cisco Nexus 9000 Series Switches xii

Documentation Feedback xii

Communications, services, and additional information xii

Cisco Bug Search Tool xiii

Documentation feedback xiii

CHAPTER 1

New and Changed Information 1

New and Changed Information 1

CHAPTER 2

Overview 3

Licensing requirements 3

Supported platforms 3

Information about Multicast 4

Multicast Distribution Trees 4

Source Trees 4

Shared Trees 5

Bidirectional Shared Trees 6

Multicast Forwarding 6

Cisco NX-OS PIM 7

ASM 8

Bidir 9

SSM 9

RPF Routes for Multicast 9

IGMP 9
IGMP Snooping 9
Interdomain Multicast 9
SSM 10
MSDP 10
MBGP 10
MRIB 10
Virtual Port Channels and Multicast 11
Guidelines and Limitations for Multicast 11
High-Availability Requirements for Multicast 12
Virtual Device Contexts 12
Troubleshooting Inconsistency Between SW and HW Multicast Routes 12

CHAPTER 3 Configuring IGMP 13
About IGMP 13
IGMP Versions 13
IGMP Basics 14
Prerequisites for IGMP 16
Guidelines and Limitations for IGMP 16
Default Settings for IGMP 17
Configuring IGMP Parameters 17
Configuring IGMP Interface Parameters 17
Configuring an IGMP SSM Translation 23
Configuring the Enforce Router Alert Option Check 24
Configuring IGMP Host Proxy 25
Overview of IGMP Host Proxy 25
IGMP Join Process 25
IGMP Leave Process 25
Guidelines and Limitations for IGMP 26
How to Configure IGMP Host Proxy 26
Restarting the IGMP Process 28
Verifying the IGMP Configuration 29
Configuration Examples for IGMP 29

CHAPTER 4 **Configuring MLD** 31

About MLD 31

MLD Versions 31

MLD Basics 32

MLD Snooping 34

Prerequisites for MLD 34

Guidelines and Limitations for MLD 34

Default Settings for MLD 35

Configuring MLD Snooping 36

Configuring MLD Parameters 39

Configuring MLD Interface Parameters 39

Configuring an MLD SSM Translation 44

Verifying the MLD Configuration 45

Verifying the MLD Snooping Configuration 46

Configuration Example for MLD 46

CHAPTER 5 **Configuring PIM and PIM6** 49

About PIM 49

PIM SSM with vPC 50

Hello Messages 51

Join-Prune Messages 51

State Refreshes 52

Rendezvous Points 52

Static RP 52

BSRs 52

Auto-RP 53

Multiple RPs Configured in a PIM Domain 54

Anycast-RP 54

PIM Register Messages 55

Designated Routers 55

Designated Forwarders 55

ASM Switchover from Shared Tree to Source Tree 56

Administratively Scoped IP Multicast 56

Multicast Counters 56
Multicast Heavy Template 56
Multicast VRF-Lite Route Leaking 57
PIM Graceful Restart 57
Generation IDs 57
PIM Graceful Restart Operations 57
PIM Graceful Restart and Multicast Traffic Flow 59
High Availability 59
Prerequisites for PIM 59
Guidelines and Limitations for PIM and PIM6 59
Guidelines and Limitations for Hello Messages 63
Guidelines and Limitations for Rendezvous Points 63
Guidelines and Limitations for Multicast VRF-lite Route Leaking 64
Default Settings 64
Configuring PIM 65
PIM Configuration Tasks 65
Enabling the PIM Feature 66
Configuring PIM Sparse Mode Parameters 66
Configuring PIM Sparse Mode Parameters 68
Configuring PIM6 Sparse Mode Parameters 71
Configuring ASM 73
Configuring Static RPs 73
Configuring BSRs 76
Configuring Auto-RP 78
Configuring a PIM Anycast-RP Set 81
Configuring Shared Trees Only for ASM 86
Configuring SSM 88
Configuring PIM SSM Over a vPC 90
Configuring RPF Routes for Multicast 91
Configuring Multicast Multipath 92
Configuring Multicast VRF-Lite Route Leaking 93
Configuring Route Maps to Control RP Information Distribution 94
Configuring Route Maps to Control RP Information Distribution (PIM) 95
Configuring Route Maps to Control RP Information Distribution (PIM6) 96

Configuring Message Filtering 97
Configuring Message Filtering 98
Configuring Message Filtering (PIM6) 100
Restarting the PIM Processes 101
Restarting the PIM Process 102
Restarting the PIM6 Process 102
Configuring BFD for PIM in VRF Mode 103
Configuring BFD for PIM in Interface Mode 104
Enabling the Multicast Heavy and Extended Heavy Template 105
Verifying the PIM Configuration 107
Displaying Statistics 108
Displaying PIM Statistics 108
Clearing PIM Statistics 108
Configuring Multicast Service Reflection 109
Guidelines and Limitations for Multicast Service Reflection 109
Prerequisites 110
Configuring Multicast Service Reflection 111
Configuration Examples for Multicast Service Reflection 115
Configuration Examples for PIM 117
SSM Configuration Example 117
PIM SSM Over vPC Configuration Example 118
BSR Configuration Example 122
Auto-RP Configuration Example 123
PIM Anycast RP Configuration Example 124
Prefix-Based and Route-Map-Based Configurations 125
Output 126
Related Documents 127
Standards 127
MIBs 127

CHAPTER 6 **Configuring IGMP Snooping** 129
About IGMP Snooping 129
IGMPv1 and IGMPv2 130
IGMPv3 130

IGMP Snooping Querier 131
Virtualization Support 131
Prerequisites for IGMP Snooping 131
Guidelines and Limitations for IGMP Snooping 132
Default Settings 133
Configuring IGMP Snooping Parameters 133
Configuring Global IGMP Snooping Parameters 133
Configuring IGMP Snooping Parameters per VLAN 136
Verifying the IGMP Snooping Configuration 140
Displaying IGMP Snooping Statistics 140
Clearing IGMP Snooping Statistics 140
Configuration Examples for IGMP Snooping 140

CHAPTER 7 **Configuring MSDP** 143
About MSDP 143
SA Messages and Caching 144
MSDP Peer-RPF Forwarding 145
MSDP Mesh Groups 145
Prerequisites for MSDP 145
Default Settings 145
Configuring MSDP 146
Enabling the MSDP Feature 146
Configuring MSDP Peers 147
Configuring MSDP Peer Parameters 148
Configuring MSDP Global Parameters 151
Configuring MSDP Mesh Groups 152
Restarting the MSDP Process 153
Verifying the MSDP Configuration 154
Monitoring MSDP 155
Displaying Statistics 155
Clearing Statistics 155
Configuration Examples for MSDP 156
Related Documents 157
Standards 157

CHAPTER 8 **Configuring MVR** 159

About MVR 159

MVR Interoperation with Other Features 160

Guidelines and Limitations for MVR 160

Default MVR Settings 160

Configuring MVR 161

Configuring MVR Global Parameters 161

Configuring MVR Interfaces 162

Suppressing IGMP Query Forwarding from VLANs 164

Verifying the MVR Configuration 165

Configuration Examples for MVR 167

CHAPTER 9 **Configuring Microsoft Network Load Balancing (NLB)** 169

About Network Load Balancing (NLB) 169

Guidelines and Limitations for NLB 170

Prerequisites for Microsoft Network Load Balancing (NLB) 171

Multicast Mode 171

IGMP Multicast Mode 172

Verifying the NLB Configuration 173

APPENDIX A **IETF RFCs for IP Multicast** 175

IETF RFCs for IP Multicast 175

APPENDIX B **Configuration Limits for Cisco NX-OS Multicast** 177

Configuration Limits 177

Preface

This preface includes the following sections:

- Audience, on page xi
- Document Conventions, on page xi
- Related Documentation for Cisco Nexus 9000 Series Switches, on page xii
- Documentation Feedback, on page xii
- Communications, services, and additional information, on page xii

Audience

This publication is for network administrators who install, configure, and maintain Cisco Nexus 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.

Related Documentation for Cisco Nexus 9000 Series Switches

The entire Cisco Nexus 9000 Series switch documentation set is available at the following URL:

https://www.cisco.com/en/US/products/ps13386/tsd_products_support_series_home.html

Documentation Feedback

To provide technical feedback on this document, or to report an error or omission, please send your comments to nexus9k-docfeedback@cisco.com. We appreciate your feedback.

Communications, services, and additional information

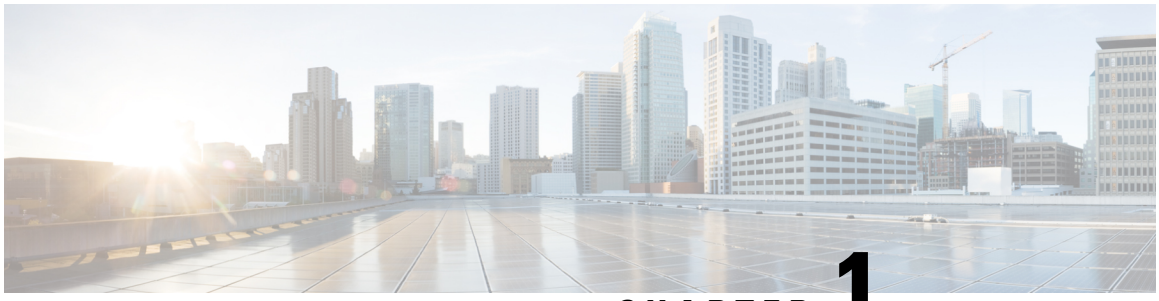
- To receive timely, relevant information from Cisco, sign up at Cisco Profile Manager.
- To get the business impact you're looking for with the technologies that matter, visit Cisco Services.
- To submit a service request, visit Cisco Support.
- To discover and browse secure, validated enterprise-class apps, products, solutions, and services, visit Cisco DevNet.
- To obtain general networking, training, and certification titles, visit Cisco Press.
- To find warranty information for a specific product or product family, access Cisco Warranty Finder.

Cisco Bug Search Tool

Cisco Bug Search Tool (BST) is a gateway to the Cisco bug-tracking system, which maintains a comprehensive list of defects and vulnerabilities in Cisco products and software. The BST provides you with detailed defect information about your products and software.

Documentation feedback

To provide feedback about Cisco technical documentation, use the feedback form available in the right pane of every online document.



CHAPTER 1

New and Changed Information

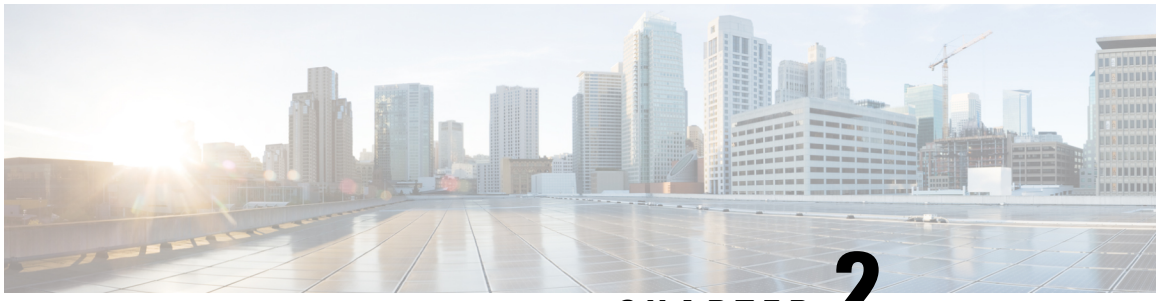
• New and Changed Information, on page 1

New and Changed Information

Table 1: New and Changed Features for Cisco NX-OS Release 10.1(x)

Feature	Description	Changed in Release	Where Documented
There are no new features for this release.		10.1(1)	

New and Changed Information



CHAPTER 2

Overview

This chapter describes the multicast features of Cisco NX-OS.

- Licensing requirements, on page 3
- Supported platforms, on page 3
- About Multicast, on page 4
- Guidelines and Limitations for Multicast, on page 11
- High-Availability Requirements for Multicast, on page 12
- Virtual Device Contexts, on page 12
- Troubleshooting Inconsistency Between SW and HW Multicast Routes , on page 12

Licensing requirements

To operate Cisco NX-OS, you must obtain and install licenses appropriate for your features and platform requirements.

- Base (Essential) and add-on licenses are available for various feature sets.
- Licenses may be permanent, temporary, or available for evaluation, depending on the product and purchase option.
- Advanced features require additional feature licenses beyond the base license.
- Licenses are applied and managed through the device CLI.

Supported platforms

The Nexus Switch Platform Support Matrix lists the following:

- Supported Cisco Nexus 9000 and 3000 switch models
- NX-OS software release versions

For the full platform-feature mapping, see the Nexus Switch Platform Support Matrix.

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.

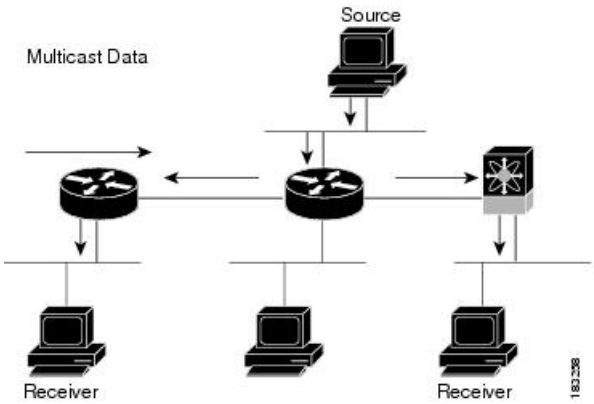
Multicast involves both a method of delivery and discovery of senders and receivers of multicast data, which is transmitted on IP multicast addresses called groups. A multicast address that includes a group and source IP address is often referred to as a channel. The Internet Assigned Number Authority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



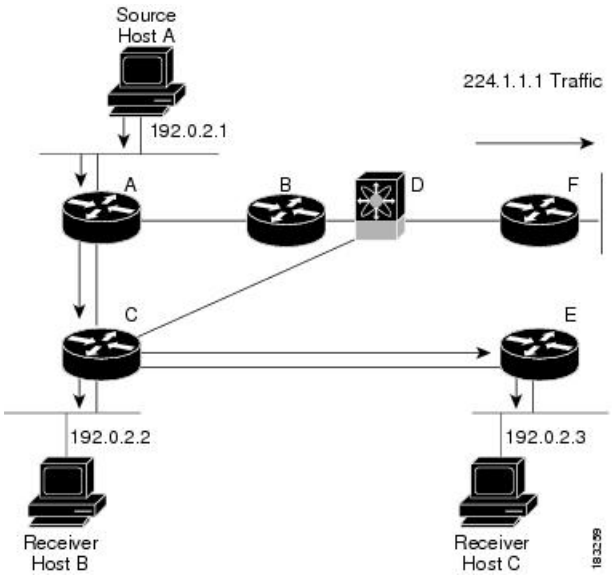
Multicast Distribution Trees

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

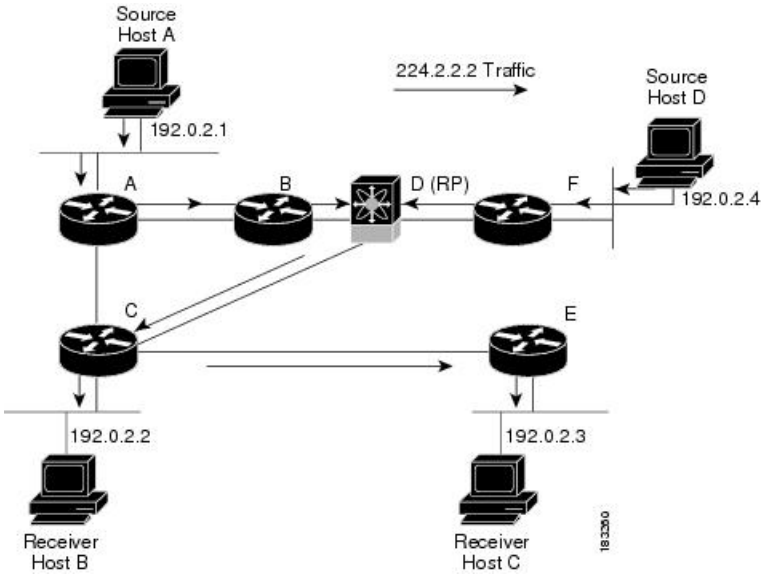


The notation (S, G) represents the multicast traffic from source S on group G. The SPT in this figure is written (192.0.2.1, 224.1.1.1). Multiple sources can be transmitting on the same group.

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 224.2.2.2 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

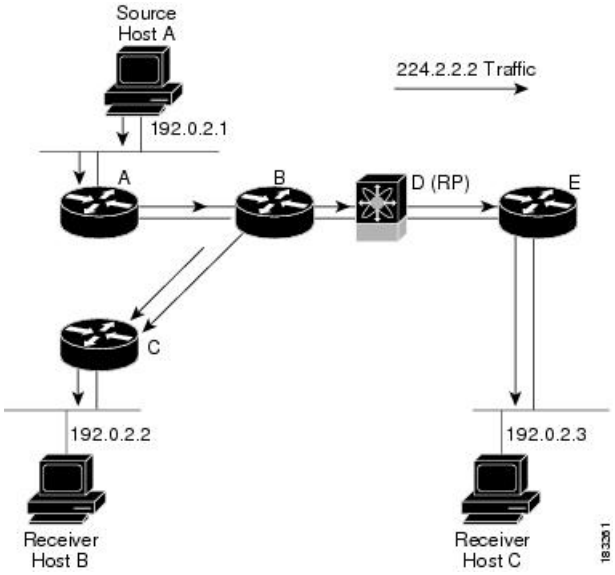


The notation (*, G) represents the multicast traffic from any source on group G. The shared tree in this figure is written (*, 224.2.2.2).

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).

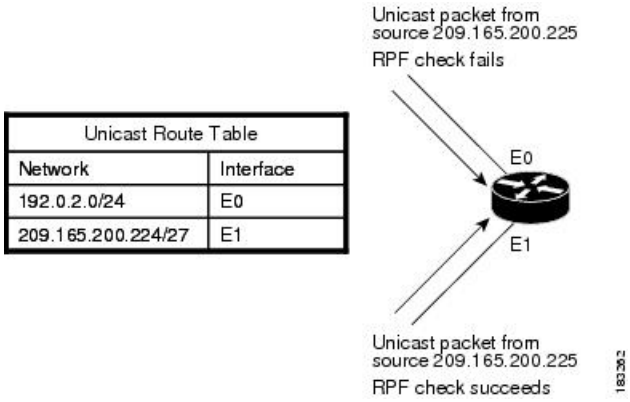
Multicast Forwarding

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 RP (ASM 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.

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



Cisco NX-OS PIM

Cisco 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 Cisco 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.

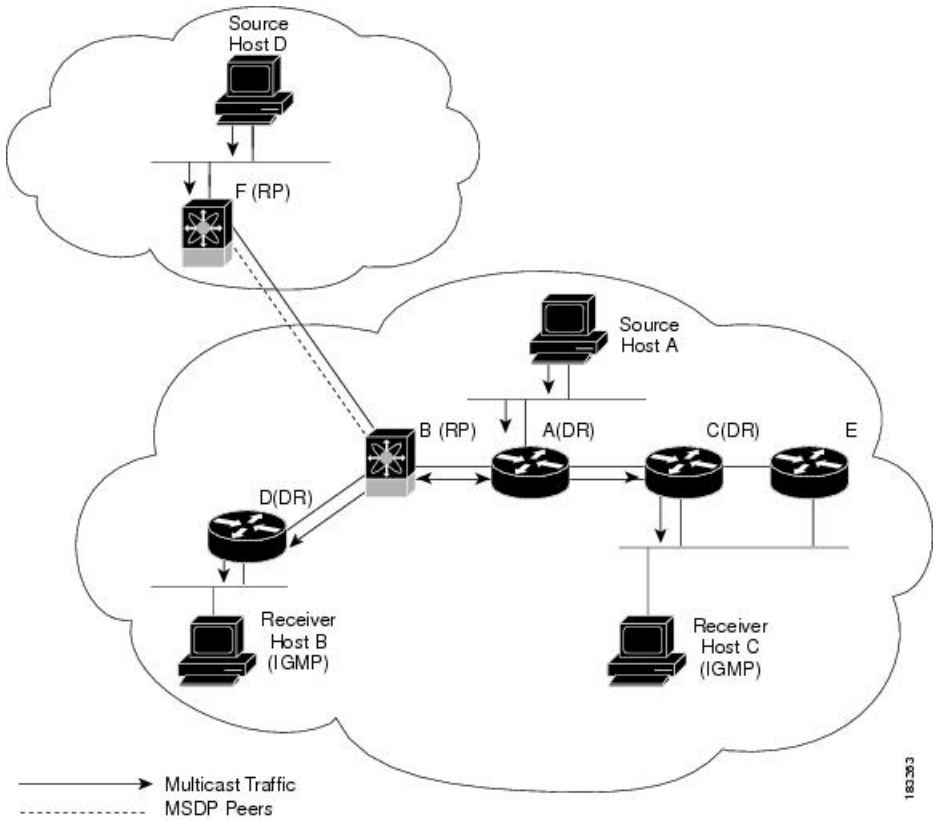
The distribution trees change automatically to reflect the topology changes due to link or router failures. PIM dynamically tracks both multicast-capable sources and receivers.

The router uses the unicast routing table and RPF routes for multicast to create multicast routing information.

Note In this publication, “PIM for IPv4” refers to the Cisco 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)

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, 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. 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

Cisco 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 Cisco 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

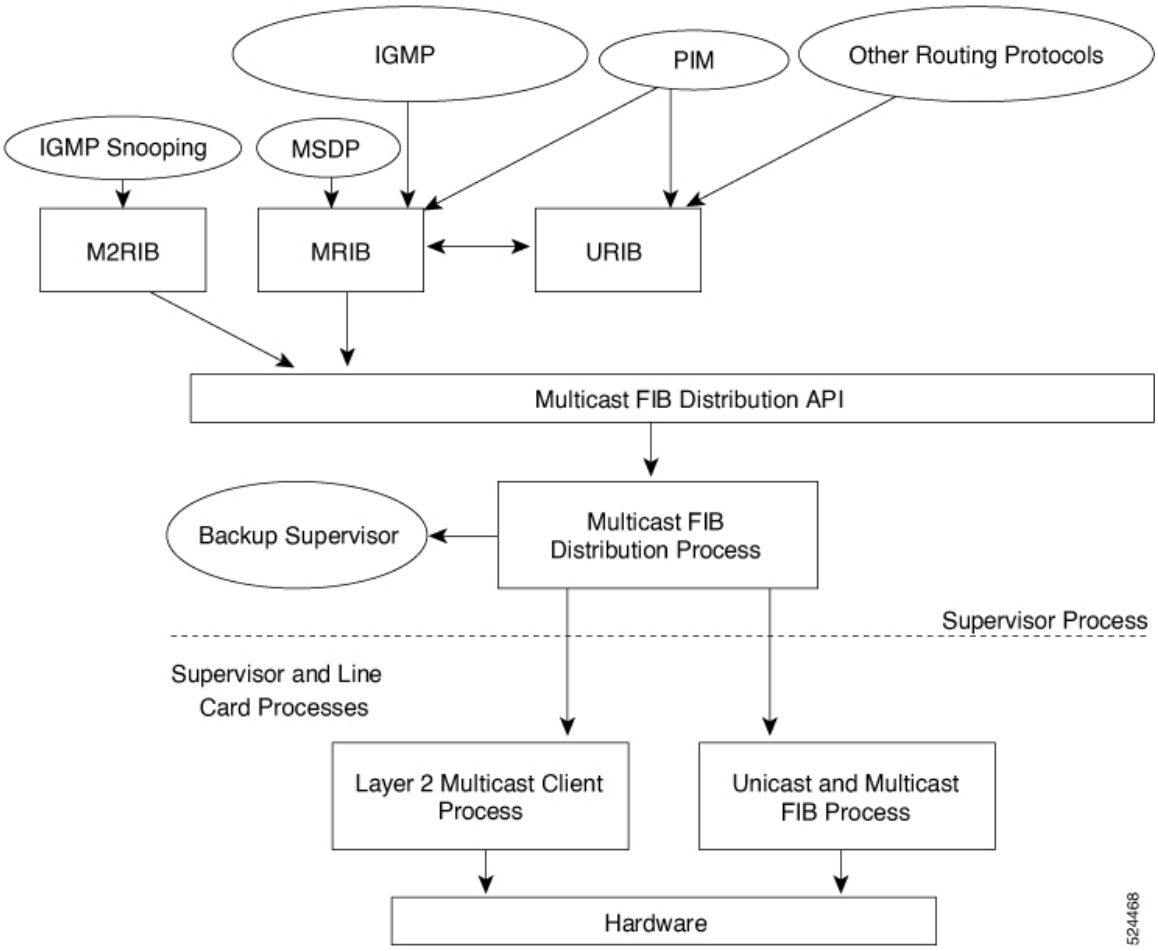
The Cisco NX-OS IPv4 Multicast Routing Information Base (MRIB) is a repository for route information that is generated by multicast protocols such as PIM and IGMP. The MRIB does not affect the route information itself. The MRIB maintains independent route information for each virtual routing and forwarding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

The major components of the Cisco 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 Cisco NX-OS multicast software architecture.

Figure 7: Cisco 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—
- IGMP snooping—You should configure the vPC peers identically.

It is recommended to configure a snooping querier on a L2 device with lower IP address to force the L2 device as the querier. This will be useful in handling the scenario where multi chassis EtherChannel trunk (MCT) is down.

Guidelines and Limitations for Multicast

- Beginning with Cisco NX-OS Release 10.1(2), Layer3 Multicast is supported on N9K-X9624-R2 line card.

- 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.
- Bidirectional mode is not supported on Cisco Nexus 9500 platform switches with -R line cards.
- IPv6 multicast is not supported on Cisco Nexus 9500 R Series line cards.

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 *Cisco Nexus 9000 Series NX-OS High Availability and Redundancy Guide*.

Virtual Device Contexts

Cisco NX-OS can segment operating system and hardware resources into virtual device contexts (VDCs) that emulate virtual devices. The Cisco Nexus 9000 Series 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

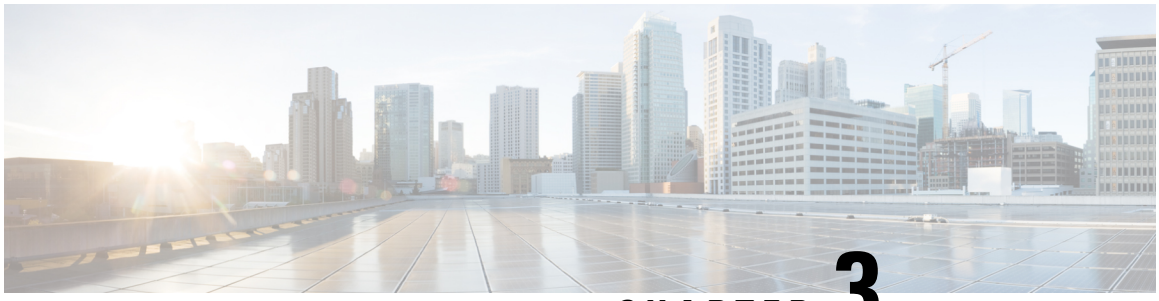
The issue can be seen when numerous active flows are received beyond the hardware capacity. This causes some of the entries not to be programmed in hardware while there is no free hardware index.

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 3

Configuring IGMP

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

- About IGMP, on page 13
- Prerequisites for IGMP, on page 16
- Guidelines and Limitations for IGMP, on page 16
- Default Settings for IGMP, on page 17
- Configuring IGMP Parameters, on page 17
- Configuring IGMP Host Proxy, on page 25
- Restarting the IGMP Process, on page 28
- Verifying the IGMP Configuration, on page 29
- Configuration Examples for IGMP, on page 29

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 Cisco Nexus 9000 Series switches do not support SSM until Cisco NX-OS Release 7.0(3)I2(1).

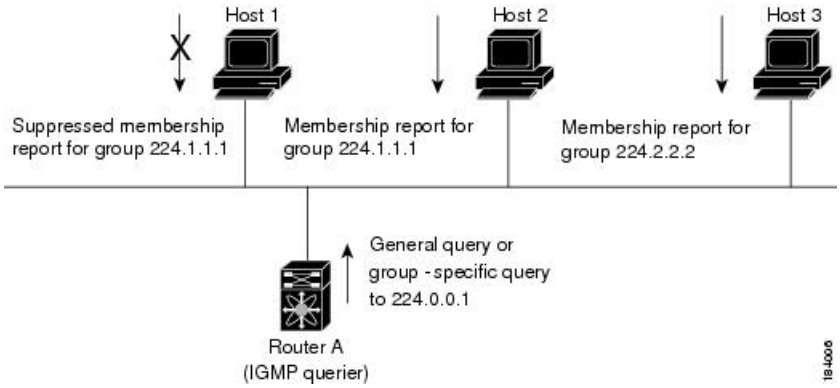
For detailed information about IGMPv2, see RFC 2236.

For detailed information about IGMPv3, see RFC 5790.

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.

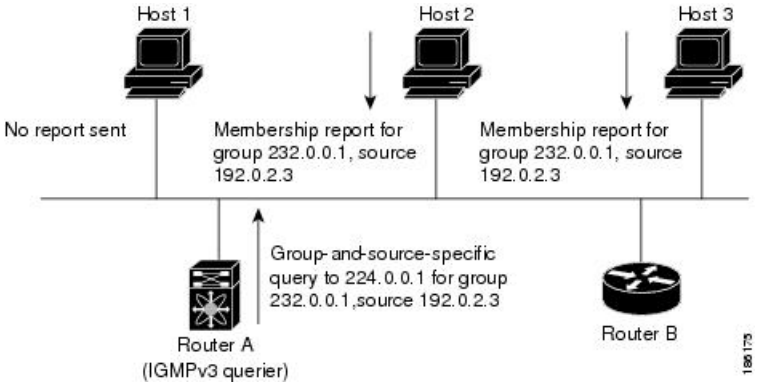
The software elects a router as the IGMP querier on a subnet if it has the lowest IP address. As long as a router continues to receive query messages from a router with a lower IP address, it resets a timer that is based on its querier timeout value. If the querier timer of a router expires, it becomes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

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.

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:

- For Cisco Nexus 9200 Series switches, the S, G routes do not expire if IGMP or source traffic originates from the same IP address.
- Sometimes vPC nodes may not have reachability to the source but may need the path to the AnycastRP pair. The multicast group states are created on the vPC peers due to the S,G,R prunes that were targeted to the RP when better route to the source is available and the traffic is coming over the shared tree.

  Since S,G is still available via the preferred static route to S, a (S,G,R) prune gets initiated to the other RP causing that state to be created. With no reachability to the source S on the VPC peers, a NULL RPF causes any traffic pulled via (*,G) to be dropped via the longest prefix match with the (S,G).

  This is a known issue. This issue can be avoided if the SPT infinity is not configured on the vPC peers or the Anycast RP reach ability from one of the RP pairs going down can be avoided to the other source via the vPC peers.
- Excluding or blocking a list of sources according to IGMPv3 (RFC 5790) is not supported.
- Beginning with Cisco NX-OS Release 9.2(2) IGMP is supported on Cisco Nexus 9500 platform switches with -R line cards.
- Configuration of nondefault IGMP related timers can be done on L3 physical interface and SVI, or in VLAN configuration mode if querier IP is configured in VLAN configuration mode. It is not recommended to configure querier IP in VLAN configuration mode if there is PIM enabled SVI for that VLAN.

  When query maximum response time (query-max-response-time) and IGMP query-interval are modified on the L3 physical interface or SVI, IGMP querier, timeout gets adjusted automatically to 2 times query interval plus MRT. To modify further, use **ip igmp querier-timeout** command for L3 physical interface.

  However, for SVI the value must be set according to the value shown in **show ip igmp interface vlan X** command output via **ip igmp snooping querier-timeout** command in VLAN configuration mode for querier election to happen as expected shell current querier become unavailable.

  For L3 physical interface, use **show ip igmp interface <intf>** command . For SVI, use **show ip igmp snooping querier <vlan>** to display relevant igmp snooping querier information. Both configuration commands should show same querier timeout for correct configuration.

  PIM hello interval determines how fast a PIM neighbor determines its peer availability. If the unavailable PIM neighbor happens to also be IGMP querier, new querier election happens at the same time as neighbor expiry (90 seconds - 3 x 30 seconds PIM hello interval). At the same time though L2 snooping querier timer dictates when new querier election is to happen (default 2 x query interval plus MRT).

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
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 Cisco IOS CLI, be aware that the Cisco NX-OS commands for this feature might differ from the Cisco 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.
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.
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.

Parameter	Description
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
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 *Cisco Nexus 9000 Series NX-OS Unicast Routing Configuration Guide*.

SUMMARY STEPS

1. **configure terminal**
2. **interface** *interface*
3. **ip igmp version** *value*
4. **ip igmp join-group** {**group** [**source** *source*] | **route-map** *policy-name*}
5. **ip igmp static-oif** {*group* [**source** *source*] | **route-map** *policy-name*}
6. **ip igmp startup-query-interval** *seconds*
7. **ip igmp startup-query-count** *count*
8. **ip igmp robustness-variable** *value*
9. **ip igmp querier-timeout** *seconds*
10. **ip igmp query-timeout** *seconds*
11. **ip igmp query-max-response-time** *seconds*
12. **ip igmp query-interval** *interval*
13. **ip igmp last-member-query-response-time** *seconds*
14. **ip igmp last-member-query-count** *count*
15. **ip igmp group-timeout** *seconds*
16. **ip igmp report-link-local-groups**
17. **ip igmp report-policy** *policy*
18. **ip igmp access-group** *policy*
19. **ip igmp immediate-leave**
20. (Optional) **show ip igmp interface** [*interface*] [**vrf** *vrf-name* | **all**] [**brief**]
21. (Optional) **copy running-config startup-config**

DETAILED STEPS

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.
		Note
		Use the commands listed from step-3 to configure the IGMP interface parameters.

Step 3	**ip igmp version** *value* **Example:** `switch(config-if)# ip igmp version 3`	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
Step 4	**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.
Step 5	**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. **Note** A source tree is built for the (S, G) state only if you enable IGMPv3.
Step 6	**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.
Step 7	**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.
Step 8	**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.
Step 9	**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.
Step 10	**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
Step 11	**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.
Step 12	**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.
Step 13	**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.
Step 14	**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.
Step 15	**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.
Step 16	**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.
Step 17	**ip igmp report-policy** *policy* **Example:** `switch(config-if)# ip igmp report-policy my_report_policy`	Configures an access policy for IGMP reports that is based on a route-map policy.
Step 18	**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.
Step 19	**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.

	Command or Action	Purpose
		Note Use this command only when there is one receiver behind the interface for a given group.
Step 20	(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 21	(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)

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

SUMMARY STEPS

1. **configure terminal**
2. **ip igmp ssm-translate** *group-prefix source-addr*
3. (Optional) **show running-configuration igmp**
4. (Optional) **copy running-config startup-config**

DETAILED STEPS

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.

SUMMARY STEPS

1. **configure terminal**
2. [**no**] **ip igmp enforce-router-alert**
3. (Optional) **show running-configuration igmp**
4. (Optional) **copy running-config startup-config**

DETAILED STEPS

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.
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 IGMP Host Proxy

This section contains the following information:

Overview of IGMP Host Proxy

The IGMP host proxy support is provided for underlay multicast on Cisco Nexus 9300 EX/FX/FX2/FX3/GX/GX2 switches with port-channel (L3) uplink. This feature is introduced in Cisco NX-OS Release 9.3(4). The IGMP host proxy feature helps to connect PIM enabled multicast network domain to a domain that does not understand PIM. This feature configures an interface as a proxy interface that proxies PIM joins/prunes that are received on the internal PIM network to IGMP joins/leaves.

IGMP Join Process

When a host wants to join a multicast group, the host sends one or more unsolicited Membership Reports for the multicast group that it wants to join. Further, IGMP joins are by default sent on receipt of an IGMP query. Unsolicited mode can be configured to periodically send the reports. Only IGMPv2 reports are sent upstream.

IGMP Leave Process

IGMPv2 leaves are sent when the last host in the multicast network leaves. Therefore on receipt of the PIM prune from the last host, IGMPv2 leaves are sent upstream to indicate no more interest.

Guidelines and Limitations for IGMP

IGMP has the following guidelines and limitations:

- The IGMP host SG proxy is not supported with vPC.
- Excluding or blocking a list of sources according to IGMPv3 (RFC 5790) is not supported.
- For Cisco Nexus 9200 Series switches, the S, G routes do not expire if IGMP or source traffic originates from the same IP address.
- IGMP is supported on Cisco Nexus 9300-FX platform switches.
- Configuring the route-map in **igmp static-oif** is limited to 255 range. When the route-map is configured with a range larger than /24 such as /8 or /4, the following log will be displayed:

```
2020 May 13 10:10:58 LO5S-NSWDDNGEF01B %IGMP-3-GROUP_RANGE_IGNORE: igmp [29534] Too
many Groups in Group Range 224.4.1.0 - 224.4.13.255
2020 May 13 12:26:13 LO5S-NSWDDNGEF01B %IGMP-3-GROUP_RANGE_IGNORE: igmp [29534] Too
many Groups in Group Range 224.4.1.0 - 224.4.13.255
2020 May 13 12:47:01 LO5S-NSWDDNGEF01B %IGMP-3-GROUP_RANGE_IGNORE: igmp [29534] Too
many Groups in Group Range 224.4.0.64 - 224.4.3.64
```

  The work around for this limitation is to split the required range to multiple 255 ranges or smaller and use the multiple route-map sequences for each range.

- Configuration of nondefault IGMP related timers can be done on L3 physical interface and SVI, or in VLAN configuration mode if querier IP is configured in VLAN configuration mode. It is not recommended to configure querier IP in VLAN configuration mode if there is PIM enabled SVI for that VLAN.

  When query maximum response time (query-max-response-time) and IGMP query-interval are modified on the L3 physical interface or SVI, IGMP querier, timeout gets adjusted automatically to 2 times query interval plus MRT. To modify further, use **ip igmp querier-timeout** command for L3 physical interface.

  However, for SVI the value must be set according to the value shown in **show ip igmp interface vlan X** command output via **ip igmp snooping querier-timeout** command in VLAN configuration mode for querier election to happen as expected shell current querier become unavailable.

  For L3 physical interface, use **show ip igmp interface <intf>** command . For SVI, use **show ip igmp snooping querier <vlan>** to display relevant igmp snooping querier information. Both configuration commands should show same querier timeout for correct configuration.

  PIM hello interval determines how fast a PIM neighbor determines its peer availability. If the unavailable PIM neighbor happens to also be IGMP querier, new querier election happens at the same time as neighbor expiry (90 seconds - 3 x 30 seconds PIM hello interval). At the same time though L2 snooping querier timer dictates when new querier election is to happen (default 2 x query interval plus MRT).

How to Configure IGMP Host Proxy

Perform the following steps to configure IGMP host proxy:

Table 6: Configuring IGMP Host Proxy

Step	Command	Purpose

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

Step 2	**interface** *interface-name* Example: `switch(config)# interface port-channel 1`	Enters interface configuration mode.
Step 3	**no shutdown** Example: `switch(config-if)# no shutdown`	Configures the interface in no shutdown mode.
Step 4	**ip address** *ip address* Example: `switch(config-if)# ip address 10.1.1.1`	Configures the IP address.
Step 5	[**no**] **ip igmp host-proxy** [**unsolicited** *time* | **route-map** *route-map-name* [**unsolicited** *time*] | **prefix-list** *prefix-list-name* [**unsolicited** *time*]] Example: `switch(config-if)# ip igmp host-proxy unsolicited 6`	Configures the IGMP host proxy for the route-map.
(Optional) **Step 6**	[**no**] **ip igmp host-proxy** [**sg-proxy**] [**unsolicited** *time* | **route-map** *route-map-name* [**unsolicited** *time*] | **prefix-list** *prefix-list-name* [**unsolicited** *time*]] Example: `switch(config-if)# ip igmp host-proxy sg-proxy 4`	Configures the IGMP SG proxy.
Step 7	**show ip igmp groups** Example: `switch(config)# show ip igmp groups`	Displays only IGMPv2 host-proxy groups (and not IGMPv3).
Step 8	**show ip igmp** *interface-name interface-number* Example: `switch(config)# show ip igmp port-channel 1`	Displays the IGMP interfaces for VRF.

Step	Command	Purpose
Step 9	**show ip igmp local-groups** *interface-name interface-number* Example: `switch(config)# show ip igmp local-groups port-channel 1`	Displays the IGMP locally joined group membership for VRF.
Step 10	**show ip pim host-proxy** Exanple: `switch(config)# show ip pim host-proxy`	Displays the PIM host proxy interfaces.

Restarting the IGMP Process

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

SUMMARY STEPS

1. **restart igmp**
2. **configure terminal**
3. **ip igmp flush-routes**
4. (Optional) **show running-configuration igmp**
5. (Optional) **copy running-config startup-config**

DETAILED STEPS

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:**	Shows the running-configuration information.

	Command or Action	Purpose
	`switch(config)# show running-configuration igmp`	
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.

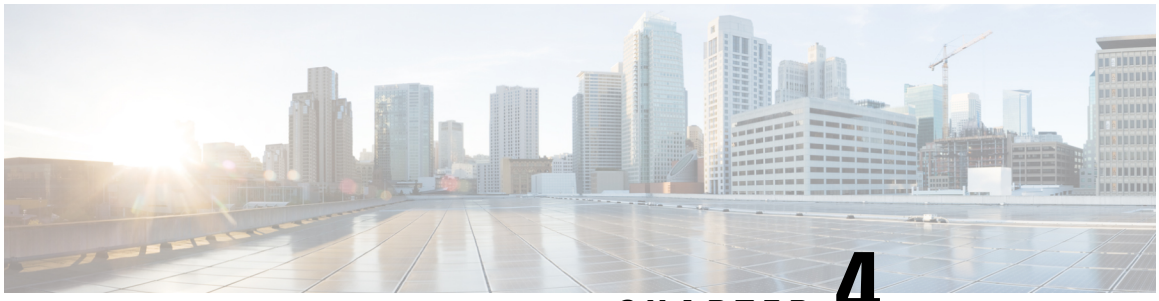
Configuration Examples for IGMP

The following example shows how to configure the IGMP parameters:

```
configure terminal

  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 4

Configuring MLD

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

- About MLD, on page 31
- Prerequisites for MLD, on page 34
- Guidelines and Limitations for MLD, on page 34
- Default Settings for MLD, on page 35
- Configuring MLD Snooping, on page 36
- Configuring MLD Parameters, on page 39
- Verifying the MLD Configuration, on page 45
- Verifying the MLD Snooping Configuration, on page 46
- Configuration Example for MLD, on page 46

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:

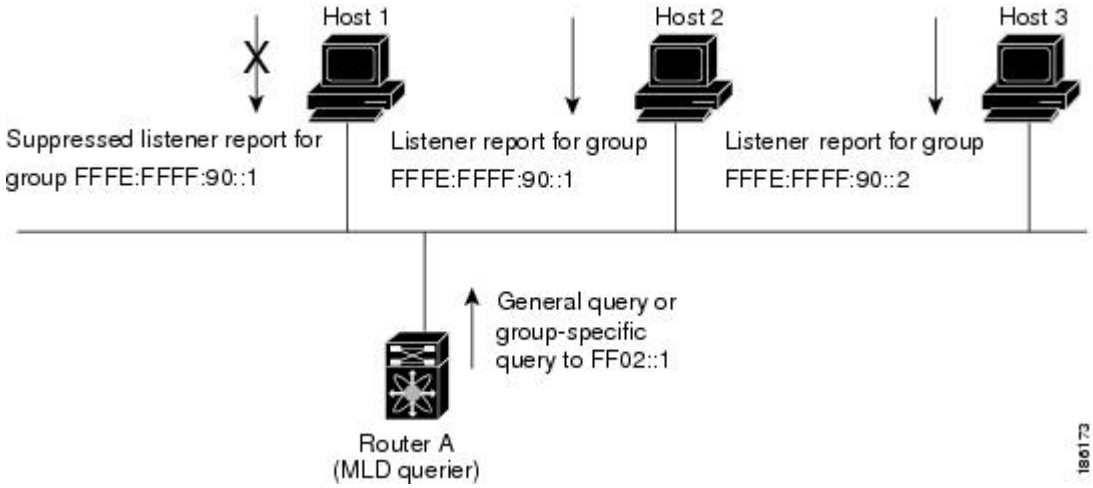
- Support for Source-Specific Multicast (SSM), which builds shortest path trees from each receiver to the source, through the following features:
  - Host messages that can specify both the group and the source.
  - The multicast state that is maintained for groups and sources, not just for groups as 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.

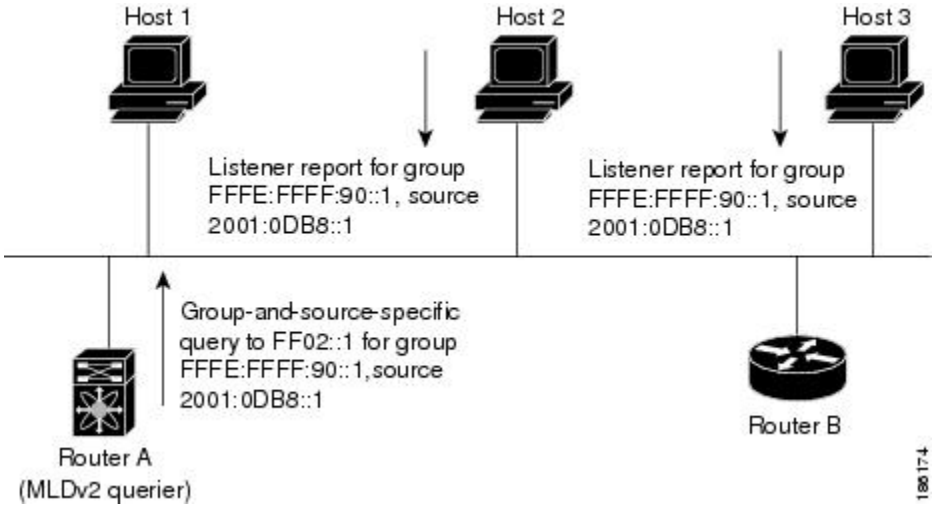
Host 1's listener report is suppressed, and host 2 sends its listener report for group FFFE:FFFF:90::1 first. Host 1 receives the report from host 2. Because only one listener report per group needs to be sent to the router, other hosts suppress their reports to reduce network traffic. Each host waits 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.

Flooding mode is used for forwarding unknown IPv6 multicast packets. In the flooding mode all endpoint groups (EPGs) and all ports under the bridge-domain will get the flooded packets.

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:

- The Cisco Nexus 9200, 9300, and 9300-EX Series switches support MLD.
- 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 aging. 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 new generation ToR switches with vPC and without vPC, which are switch models with "EX", "FX" or "FX2" at the end of the switch name; and on EoR switches with "EX" and "FX" line cards.
- Beginning with Cisco NX-OS Release 9.3(5), IPv6 MLD snooping is supported on Cisco Nexus 9500 platform switches.
- MLD snooping is also supported on the following T2 line cards in a EOR switch: N9K-X9636PQ, N9K-X9408PC-CFP2, N9K-X9432PQ, N9K-X9464PX, N9K-X9464TX, N9K-X9464TX2.
- MLD snooping is supported on all Cisco Nexus 9000 and Cisco Nexus 3000 platforms with T2, T2P, T3, TH, TH2 and T2 EORs. It is not supported on the Cisco Nexus 9000 T2 TORs — N9K-C9372PX, N9K-C9372PX-E, N9K-C9372TX, N9K-C9372TX-E, N9K-C9332PQ, N9K-C93128TX, N9K-C9396PX, N9K-C9396TX.
- MLD snooping is not supported on the FEX ports and on Network Load Balancing (NLB). It is also not supported when VLAN is in MAC mode.
- If the below commands are configured, the MLD snooping configuration will be denied at the global level:
  - ip pim cpu-punt dr-only
  - ipv6 pim cpu-punt dr-only
  - ip pim non-dr flood
  - ipv6 pim non-dr flood
- Beginning with Cisco NX-OS Release 9.3(5), MLD snooping is supported on Cisco Nexus 9300-FX3 platform switches.

Default Settings for MLD

Table 7: 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

Parameters	Default
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
Immediate leave	Disabled

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 Cisco Nexus 9000 Series platform. Both step 2 and step 3 are required to completely enable snooping on the Cisco Nexus 9000 Series platform. Reload the switch after configuring this command.
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.

	Command or Action	Purpose
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 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.

	Command or Action	Purpose
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.
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-config)# 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-config)# 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-config)# 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-config)# 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-config)# 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-config)# 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.

Configuring MLD Parameters

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

Note Before you can configure MLD snooping, enable the MLD feature using the **ipv6 mld snooping** and **system mld snooping** commands.

Configuring MLD Interface Parameters

Table 8: 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.
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.

Parameter	Description
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.
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.

Parameter	Description
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.
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 *Cisco Nexus 9000 Series 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.
		Note
		Use the commands listed from step-3 to configure the MLD interface parameters.

Step 3	**ipv6 mld version** *value* **Example:** `switch(config-if)# ipv6 mld version 2`	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.
Step 4	**ipv6 mld join-group** {**group** [**source** *source*] | **route-map** *policy-name*}	Statically binds a multicast group to the interface. If you specify only the group address, the (*, G) state is created.

	Command or Action	Purpose
	Example: `switch(config-if)# ipv6 mld join-group FFFE::1`	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. **Caution** The device CPU must handle the traffic generated by using this command.
Step 5	**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. **Note** The maximum number of groups supported per entry in the route map is 256.
Step 6	**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.
Step 7	**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.
Step 8	**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.
Step 9	**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.

	Command or Action	Purpose
Step 10	**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. **Note** This command has the same functionality as the **ipv6 mld querier-timeout** command.
Step 11	**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.
Step 12	**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.
Step 13	**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.
Step 14	**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.
Step 15	**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.
Step 16	**ipv6 mld report-link-local-groups** **Example:** `switch(config-if)# ipv6 mld 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.
Step 17	**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.
Step 18	**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**

	Command or Action	Purpose
		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.
Step 19	**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. **Note** Use this command only when there is one receiver behind the interface for a given group.
Step 20	(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 9: 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 10: Example Result of Applying SSM Translations

MLDv1 Listener Report	Resulting M6RIB Route
FF32:40::40	(2001:0DB8:0:ABCD::4, FF32:40::40)
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 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 local-groups	Displays the MLD local group membership.

The following example displays the **show ipv6 mld groups** command output. This output shows ten interfaces are sending MLD joins to group ff03:0:0:1::1 out of which nine interfaces are sending MLDv1 joins and the tenth interface is sending MLDv2 join with source 2005:0:0:1::2. There are nine entries for the group and tenth entry is appended as the source entry.

```
switch# show ipv6 mld groups vrf vrf1
MLD Connected Group Membership for VRF "VRF1" - 52 total entries
Type: S - Static, D - Dynamic, L - Local, T - SSM Translated, H - Host Proxy
      * - Cache Only
Group Address      Type Interface              Uptime    Expires   Last Reporter
ff03:0:0:1::1      D   Ethernet3/25.1         00:02:13  00:03:47  fe80::1
ff03:0:0:1::1      D   Ethernet3/25.3         00:02:13  00:04:12  fe80::2:0:0:1
ff03:0:0:1::1      D   Ethernet3/25.5         00:02:13  00:02:26  fe80::4:0:0:1
ff03:0:0:1::1      D   Ethernet3/25.4         00:02:13  00:03:31  fe80::3:0:0:1
ff03:0:0:1::1      D   Ethernet3/25.6         00:02:13  00:02:47  fe80::5:0:0:1
```

```
ff03:0:0:1::1      D   Ethernet3/25.7         00:02:13  00:03:10  fe80::6:0:0:1
ff03:0:0:1::1      D   Ethernet3/25.8         00:02:13  00:03:56  fe80::7:0:0:1
ff03:0:0:1::1      D   Ethernet3/25.9         00:02:13  00:03:28  fe80::8:0:0:1
  2005:0:0:1::2    D    Ethernet3/25.10        2d15h     00:03:37  fe80::9:0:0:1
```

Verifying the MLD Snooping Configuration

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

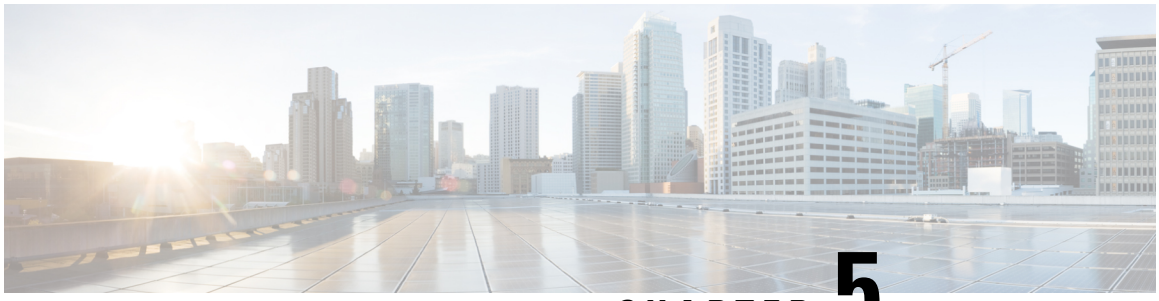
show ipv6 mld snooping [**vlan** *vlan-id*]	Displays the MLD snooping status and details for a given VLAN or all VLANs.
show ipv6 mld snooping mrouter [**vlan** *vlan-id*]	Displays the multicast router ports in each VLAN.
show ipv6 mld snooping querier [**vlan** *vlan-id*]	Displays details on the MLD 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 5

Configuring PIM and PIM6

This chapter describes how to configure the Protocol Independent Multicast (PIM) and PIM6 features on Cisco NX-OS devices in your IPv4 and IPv6 networks.

- About PIM, on page 49
- Prerequisites for PIM, on page 59
- Guidelines and Limitations for PIM and PIM6, on page 59
- Default Settings, on page 64
- Configuring PIM, on page 65
- Verifying the PIM Configuration, on page 107
- Displaying Statistics, on page 108
- Configuring Multicast Service Reflection, on page 109
- Configuration Examples for PIM, on page 117
- Related Documents, on page 127
- Standards, on page 127
- MIBs, on page 127

About PIM

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.

Cisco NX-OS supports PIM sparse mode for IPv4 networks (PIM). In PIM sparse mode, multicast traffic is sent only to locations of the network that specifically request it. You can configure PIM to run simultaneously on a router. You can use PIM global parameters to configure rendezvous points (RPs), message packet filtering, and statistics. You can use PIM interface parameters to enable multicast, identify PIM borders, set the PIM hello message interval, and set the designated router (DR) priority.

Note Cisco NX-OS does not support PIM dense mode.

In Cisco NX-OS, multicast is enabled only after you enable the PIM feature on each router and then enable PIM sparse mode on each interface that you want to participate in multicast. You can configure PIM for an

IPv4 network . In an IPv4 network, if you have not already enabled IGMP on the router, PIM enables it automatically.

You use the PIM global configuration parameters to configure the range of multicast group addresses to be handled by these distribution modes:

- Any Source Multicast (ASM) provides discovery of multicast sources. It builds a shared tree between sources and receivers of a multicast group and supports switching over to a source tree when a new receiver is added to a group. ASM mode requires that you configure an RP.

For more information about PIM sparse mode and shared distribution trees used by the ASM mode, see RFC 4601.

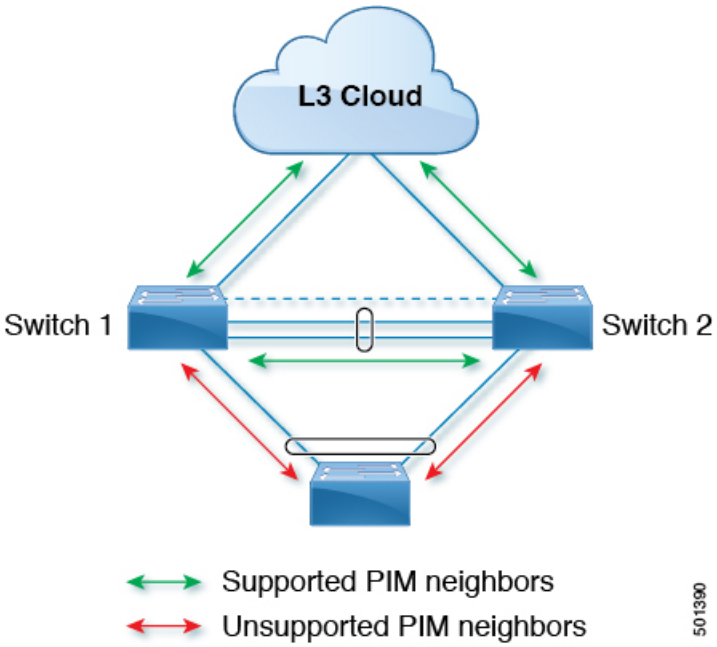
PIM SSM with vPC

Beginning with Cisco NX-OS Release 7.0(3)I4(1), you can enable PIM SSM on Cisco Nexus 9000 Series switches with an upstream Layer 3 cloud along with the vPC feature.

A PIM adjacency between a Switched Virtual Interface (SVI) on a vPC VLAN (a VLAN that is carried on a vPC Peer-Link) and a downstream device is not supported; this configuration can result in dropped multicast packets. If a PIM neighbor relationship is required with a downstream device, a physical Layer 3 interface must be used on the Nexus switches instead of a vPC SVI.

For SVIs on vPC VLANs, only one PIM adjacency is supported, which is with the vPC peer switch. PIM adjacencies over the vPC peer-link with devices other than the vPC peer switch for the vPC-SVI are not supported.

Note Cisco Nexus 9508 switches with the N9K-X9636C-R and N9K-X9636Q-R line cards support PIM SSM beginning with Cisco NX-OS Release 7.0(3)F2(1) but do not support PIM SSM on vPCs until Cisco NX-OS Release 7.0(3)F3(1). The N9K-X9636C-RX line card supports PIM SSM with and without vPCs beginning with Cisco NX-OS Release 7.0(3)F3(1).



Hello Messages

The PIM process begins when the router establishes PIM neighbor adjacencies by sending PIM hello messages to the multicast IPv4 address 224.0.0.13 or IPv6 address ff02::d. Hello messages are sent periodically at the interval of 30 seconds. When all neighbors have replied, the PIM software chooses the router with the highest priority in each LAN segment as the designated router (DR). The DR priority is based on a DR priority value in the PIM hello message. If the DR priority value is not supplied by all routers, or the priorities match, the highest IP address is used to elect the DR.

The hello message also contains a hold-time value, which is typically 3.5 times the hello interval. If this hold time expires without a subsequent hello message from its neighbor, the device detects a PIM failure on that link.

The configured hold-time changes may not take effect on first two hellos sent after enabling or disabling PIM on an interface. For the first two hellos sent on the interface, thereafter, the configured hold times will be used. This may cause the PIM neighbor to set the incorrect neighbor timeout value for the initial neighbor setup until a hello with the correct hold time is received.

For added security, you can configure an MD5 hash value that the PIM software uses to authenticate PIM hello messages with PIM neighbors.

Join-Prune Messages

When the DR receives an IGMP membership report message from a receiver for a new group or source, the DR creates a tree to connect the receiver to the source by sending a PIM join message out the interface toward the rendezvous point (ASM mode). The rendezvous point (RP) is the root of a shared tree, which is used by all sources and hosts in the PIM domain in the ASM mode.

When the DR determines that the last host has left a group or source, it sends a PIM prune message to remove the path from the distribution tree.

The routers forward the join or prune action hop by hop up the multicast distribution tree to create (join) or tear down (prune) the path.

Note In this publication, the terms “PIM join message” and “PIM prune message” are used to simplify the action taken when referring to the PIM join-prune message with only a join or prune action.

Join-prune messages are sent as quickly as possible by the software. You can filter the join-prune messages by defining a routing policy.

State Refreshes

PIM requires that multicast entries are refreshed within a 3.5-minute timeout interval. The state refresh ensures that traffic is delivered only to active listeners, and it keeps routers from using unnecessary resources.

Due to scalability optimizations in NX-OS, it can take up to 8 minutes for an OIF entry to be deleted after Join packets for the entry are no longer received, if there is still active traffic for the group or other OIFs are still present.

To maintain the PIM state, the last-hop DR sends join-prune messages once per minute. State creation applies to both (*, G) and (S, G) states as follows:

- (*, G) state creation example—An IGMP (*, G) report triggers the DR to send a (*, G) PIM join message toward the RP.
- (S, G) state creation example—An IGMP (S, G) report triggers the DR to send an (S, G) PIM join message toward the source.

If the state is not refreshed, the PIM software tears down the distribution tree by removing the forwarding paths in the multicast outgoing interface list of the upstream routers.

Rendezvous Points

A rendezvous point (RP) is a router that you select in a multicast network domain that acts as a shared root for a multicast shared tree. You can configure as many RPs as you like, and you can configure them to cover different group ranges.

Static RP

You can statically configure an RP for a multicast group range. You must configure the address of the RP on every router in the domain.

You can define static RPs for the following reasons:

- To configure routers with the Anycast-RP address
- To manually configure an RP on a device

BSRs

The bootstrap router (BSR) ensures that all routers in the PIM domain have the same RP cache as the BSR. You can configure the BSR to help you select an RP set from BSR candidate RPs. The function of the BSR

is to broadcast the RP set to all routers in the domain. You select one or more candidate BSRs to manage the RPs in the domain. Only one candidate BSR is elected as the BSR for the domain.

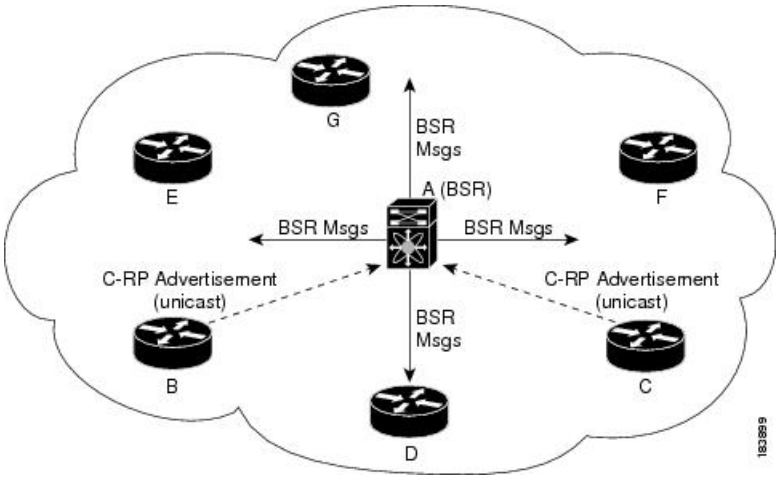
BSR is supported on Cisco Nexus 9300-FX, Cisco Nexus 9300-FX2, and Cisco Nexus 9300-FX3S platform switches.

BSR is supported on Cisco Nexus 9300-EX/FX/FX2/FX3/GX/C and 9500-EX/FX/GX platform switches.

This figure shows the BSR mechanism. Router A, the software-elected BSR, sends BSR messages out all enabled interfaces (shown by the solid lines in the figure). The messages, which contain the RP set, are flooded hop by hop to all routers in the network. Routers B and C are candidate RPs that send their candidate-RP advertisements directly to the elected BSR (shown by the dashed lines in the figure).

The elected BSR receives candidate-RP messages from all the candidate RPs in the domain. The bootstrap message sent by the BSR includes information about all of the candidate RPs. Each router uses a common algorithm to select the same RP address for a given multicast group.

Figure 12: BSR Mechanism



In the RP selection process, the RP address with the best priority is determined by the software. If the priorities match for two or more RP addresses, the software might use the RP hash in the selection process. Only one RP address is assigned to a group.

By default, routers are not enabled to listen or forward BSR messages. You must enable the BSR listening and forwarding feature so that the BSR mechanism can dynamically inform all routers in the PIM domain of the RP set assigned to multicast group ranges.

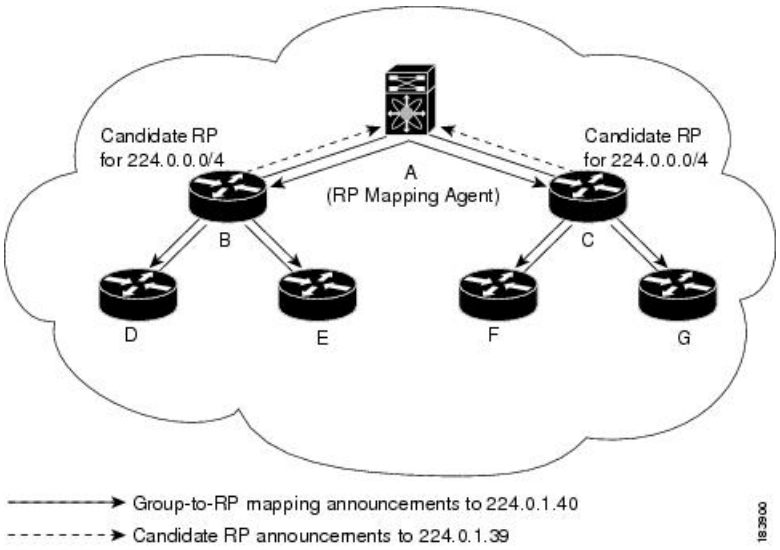
Note The BSR mechanism is a nonproprietary method of defining RPs that can be used with third-party routers.

Auto-RP

Auto-RP is a Cisco protocol that was introduced prior to the Internet standard bootstrap router mechanism. You configure Auto-RP by selecting candidate mapping agents and RPs. Candidate RPs send their supported group range in RP-Announce messages to the Cisco RP-Announce multicast group 224.0.1.39. An Auto-RP mapping agent listens for RP-Announce messages from candidate RPs and forms a Group-to-RP mapping table. The mapping agent multicasts the Group-to-RP mapping table in RP-Discovery messages to the Cisco RP-Discovery multicast group 224.0.1.40.

This figure shows the Auto-RP mechanism. Periodically, the RP mapping agent multicasts the RP information that it receives to the Cisco-RP-Discovery group 224.0.1.40 (shown by the solid lines in the figure).

Figure 13: Auto-RP Mechanism



By default, routers are not enabled to listen or forward Auto-RP messages. You must enable the Auto-RP listening and forwarding feature so that the Auto-RP mechanism can dynamically inform routers in the PIM domain of the group-to-RP mapping.

⚠

Caution Do not configure both Auto-RP and BSR protocols in the same network.

Multiple RPs Configured in a PIM Domain

This section describes the election process rules when multiple RPs are configured in a PIM domain.

Anycast-RP

Anycast-RP has two implementations: one uses Multicast Source Discovery Protocol (MSDP) and the other is based on *RFC 4610, Anycast-RP Using Protocol Independent Multicast (PIM)*. This section describes how to configure PIM Anycast-RP.

You can use PIM Anycast-RP to assign a group of routers, called the Anycast-RP set, to a single RP address that is configured on multiple routers. The set of routers that you configure as Anycast-RPs is called the Anycast-RP set. This method is the only RP method that supports more than one RP per multicast group, which allows you to load balance across all RPs in the set. The Anycast RP supports all multicast groups.

PIM register messages are sent to the closest RP, and PIM join-prune messages are sent in the direction of the closest RP as determined by the unicast routing protocols. If one of the RPs goes down, unicast routing ensures these messages will be sent in the direction of the next-closest RP.

You must configure PIM on the loopback interface that is used for the PIM Anycast RP and the PIM Bidir RP.

For more information about PIM Anycast-RP, see RFC 4610.

PIM Register Messages

PIM register messages are unicast to the RP by designated routers (DRs) that are directly connected to multicast sources. The PIM register message has the following functions:

- To notify the RP that a source is actively sending to a multicast group.
- To deliver multicast packets sent by the source to the RP for delivery down the shared tree.

The DR continues to send PIM register messages to the RP until it receives a Register-Stop message from the RP. The RP sends a Register-Stop message in either of the following cases:

- The RP has no receivers for the multicast group being transmitted.
- The RP has joined the SPT to the source but has not started receiving traffic from the source.

The PIM triggered register is enabled by default.

You can use the **ip pim register-source** command to configure the IP source address of register messages when the IP source address of a register message is not a uniquely routed address to which the RP can send packets. This situation might occur if the source address is filtered so that the packets sent to it are not forwarded or if the source address is not unique to the network. In these cases, the replies sent from the RP to the source address will fail to reach the DR, resulting in Protocol Independent Multicast sparse mode (PIM-SM) protocol failures.

The following example shows how to configure the IP source address of the register message to the loopback 3 interface of a DR:

```
ip pim register-source loopback 3
```

Note In Cisco NX-OS, PIM register messages are rate limited to avoid overwhelming the RP.

You can filter PIM register messages by defining a routing policy.

Designated Routers

In PIM ASM mode, the software chooses a designated router (DR) from the routers on each network segment. The DR is responsible for forwarding multicast data for specified groups and sources on that segment.

The DR for each LAN segment is determined as described in the Hello messages.

In ASM mode, the DR is responsible for unicasting PIM register packets to the RP. When a DR receives an IGMP membership report from a directly connected receiver, the shortest path is formed to the RP, which may or may not go through the DR. The result is a shared tree that connects all sources transmitting on the same multicast group to all receivers of that group.

Designated Forwarders

In PIM Bidir mode, the software chooses a designated forwarder (DF) at RP discovery time from the routers on each network segment. The DF is responsible for forwarding multicast data for specified groups on that segment. The DF is elected based on the best metric from the network segment to the RP.

If the router receives a packet on the RPF interface toward the RP, the router forwards the packet out all interfaces in the OIF-list. If a router receives a packet on an interface on which the router is the elected DF for that LAN segment, the packet is forwarded out all interfaces in the OIF-list except the interface that it was received on and also out the RPF interface toward the RP.

Note Cisco NX-OS puts the RPF interface into the OIF-list of the MRIB but not in the OIF-list of the MFIB.

ASM Switchover from Shared Tree to Source Tree

Note Cisco NX-OS puts the RPF interface into the OIF-list of the MRIB but not into the OIF-list of the MFIB.

In ASM mode, the DR that is connected to a receiver switches over from the shared tree to the shortest-path tree (SPT) to a source unless you configure the PIM parameter to use shared trees only.

During the switchover, messages on the SPT and shared tree might overlap. These messages are different. The shared tree messages are propagated upstream toward the RP, while SPT messages go toward the source.

For information about SPT switchovers, see the “Last-Hop Switchover to the SPT" section in RFC 4601.

Administratively Scoped IP Multicast

The administratively scoped IP multicast method allows you to set boundaries on the delivery of multicast data. For more information, see RFC 2365.

You can configure an interface as a PIM boundary so that PIM messages are not sent out on that interface.

You can use the Auto-RP scope parameter to set a time-to-live (TTL) value.

Multicast Counters

Multicast flow counters collection can be enabled in two different ways.

- Enable multicast heavy template as described in the Enabling the Multicast Heavy Template section.
- Configure the **hardware profile multicast flex-stats-enable** command in the default template.

Only Cisco Nexus 9300-EX, X9700-FX, 9300-FX, and 9300-FX2 Series switches support multicast counters. These counters provide more granularity and visibility about multicast traffic. Specifically, they show an absolute multicast packet count (bytes and rate for every multicast S,G route). These counters are valid only for S,G routes and not for *,G routes. Multicast counters appear in the output of the **show ip mroute detail**and **show ip mroute summary** commands when the multicast heavy template is enabled.

Multicast Heavy Template

You can enable the multicast heavy template in order to support significantly more multicast routes and to display multicast counters in the output of the **show ip mroute** command.

The multicast heavy template is supported for the following devices and releases:

- Cisco Nexus N9K-X9732C-EX, N9K-X9736C-E, and N9K-X97160YC-EX line cards, beginning with Cisco NX-OS Release 7.0(3)I3(2), but only for increased scalability
- Cisco Nexus 9300-EX Series switches, beginning with Cisco NX-OS Release 7.0(3)I6(1), for both increased scalability and multicast counters
- Cisco Nexus 9300-FX Series switches, beginning with Cisco NX-OS Release 7.0(3)I7(1), for both increased scalability and multicast counters

Multicast VRF-Lite Route Leaking

Beginning with Cisco NX-OS Release 7.0(3)I7(1), multicast receivers can forward IPv4 traffic across VRFs. In previous releases, multicast traffic can flow only within the same VRF.

With multicast VRF-lite route leaking, Reverse Path Forwarding (RPF) lookup for multicast routes in the receiver VRF can be performed in the source VRF. Therefore, traffic originating from the source VRF can be forwarded to the receiver VRF.

PIM Graceful Restart

Protocol Independent Multicast (PIM) graceful restart is a multicast high availability (HA) enhancement that improves the convergence of multicast routes (mroutes) after a route processor (RP) switchover. In the event of an RP switchover, the PIM graceful restart feature utilizes the generation ID (GenID) value (defined in RFC 4601) as a mechanism to trigger adjacent PIM neighbors on an interface to send PIM join messages for all (*, G) and (S, G) states that use that interface as a reverse path forwarding (RPF) interface. This mechanism enables PIM neighbors to immediately reestablish those states on the newly active RP.

Generation IDs

A generation ID (GenID) is a randomly generated 32-bit value that is regenerated each time Protocol Independent Multicast (PIM) forwarding is started or restarted on an interface. In order to process the GenID value in PIM hello messages, PIM neighbors must be running Cisco software with an implementation of PIM that is compliant with RFC 4601.

Note PIM neighbors that are not compliant with RFC 4601 and are unable to process GenID differences in PIM hello messages will ignore the GenIDs.

PIM Graceful Restart Operations

This figure illustrates the operations that occur after a route processor (RP) switchover on devices that support the PIM graceful restart feature.

Figure 14: PIM Graceful Restart Operations During an RP Switchover

The PIM graceful restart operations are as follows:

- In steady state, PIM neighbors exchange periodic PIM hello messages.
- An active RP receives PIM joins periodically to refresh multicast route (mroute) states.
- When an active RP fails, the standby RP takes over to become the new active RP.
- The new active RP then modifies the generation ID (GenID) value and sends the new GenID in PIM hello messages to adjacent PIM neighbors.
- Adjacent PIM neighbors that receive PIM hello messages on an interface with a new GenID send PIM graceful restart for all (*, G) and (S, G) mroutes that use that interface as an RPF interface.
- Those mroute states are then immediately reestablished on the newly active RP.

PIM Graceful Restart and Multicast Traffic Flow

Multicast traffic flow on PIM neighbors is not affected if the multicast traffic detects support for PIM graceful restart PIM or PIM hello messages from a node with the failing RP within the default PIM hello hold-time interval. Multicast traffic flow on a failing RP is not affected if it is non-stop forwarding (NSF) capable.

Caution The default PIM hello hold-time interval is 3.5 times the PIM hello period. Multicast high availability (HA) operations might not function as per design if you configure the PIM hello interval with a value lower than the default value of 30 seconds.

High Availability

When a route processor reloads, multicast traffic across VRFs behaves the same as traffic forwarded within the same VRF.

For information about high availability, see the *Cisco Nexus 9000 Series NX-OS High Availability and Redundancy Guide*.

Prerequisites for PIM

- 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 PIM and PIM6

PIM and PIM6 have the following guidelines and limitations:

- Cisco NX-OS PIM and PIM6 are supported on Cisco Nexus 9300-EX, Cisco Nexus 9300-FX, Cisco Nexus 9300-FX2, and Cisco Nexus 9300-FX3S platform switches.
- Configuring a secondary IP address as an RP address is not supported.

- For most Cisco Nexus devices, RPF failure traffic is dropped and sent to the CPU at a very low rate to trigger PIM asserts. For the Cisco Nexus 9000 Series switches, RPF failure traffic is always copied to the CPU in order to learn multicast sources.
- For first-hop source detection in most Cisco Nexus devices, traffic coming from the first hop is detected based on the source subnet check, and multicast packets are copied to the CPU only if the source belongs to the local subnet. The Cisco Nexus 9000 Series switches cannot detect the local source, so multicast packets are sent to the supervisor to learn the local multicast source.
- Cisco NX-OS PIM and PIM6 do not interoperate with any version of PIM dense mode or PIM Sparse Mode version 1.
- PIM SSM and PIM ASM is supported on all Cisco Nexus 9000 Series switches.
- Cisco Nexus 9000 Series switches support PIM SSM on vPCs.
- It is recommended to configure a snooping querier on a L2 device with lower IP address to force the L2 device as the querier. This will be useful in handling the scenario where multi chassis EtherChannel trunk (MCT) is down.
- When the Rendezvous Point receives a PIM Data Register, it is expected for the register to be punted up to the CPU for processing. During this operation, the register will be decapsulated and the data portion of it will be software forwarded if there are any relevant OIFs for the group.
- (S,G) route is created on the control plane when data packet is punted regardless of spt-threshold infinity configuration. The route is created to protect the control plane and limit further punting of packets.
- If the NAT flows are established before the service interface is created as shown below, use the **clear ip mroute** *group source* command to manually clear the affected routes:

```
2024 Jan 30 15:26:17.127933 MFX2-4
%IPFIB-SLOT1-2-MFIB_EGR_NAT_INVALID_INTF: Service Intf Ethernet1/31.100
not available, Impacted translation flow:
(118.4.0.1,2.1.13.153)->(228.4.11.49,204.0.1.59)L4(0,0)2024 Jan 30
15:26:23.039119 MFX2-4 %ETHPORT-5-IF_UP: Interface Ethernet1/31.100
is up in Layer3
```

- Beginning with Cisco NX-OS Release 9.2(3):
  - PIM6 on TOR is supported in multicast heavy, ext-heavy, and default templates.
  - PIM6 on the Cisco Nexus 9500 boxes with EX/FX/GX line cards is only supported in multicast heavy, ext-heavy, dual-stack-multicast templates.
- Beginning with Cisco NX-OS Release 9.3(3), PIM6 support for SVI is introduced on TOR with or without vPC for switches ending with "EX", "FX", "FX2" and on EOR for switches ending with "EX", "FX".
- PIM6 support on SVI is possible only after the MLD snooping is enabled.
- Beginning with Cisco NX-OS Release 9.3(5), PIM6 support for SVI is introduced on Cisco Nexus 9300-GX platform switches and Cisco Nexus 9500 platform switches.
- Cisco Nexus 9000 Series switches support PIM ASM and SSM on vPCs.
- Cisco Nexus 9000 Series switches do not support PIM adjacency with a vPC leg or with a router behind a vPC.
- PIM Snooping is not supported on Cisco Nexus 9000 Series switches.

- Cisco Nexus 9000 Series switches support PIM6 ASM and SSM.

Note Only Cisco Nexus 9500 Series switches with N9K-X9400 or N9K-X9500 line cards and/or N9K-C9504-FM, N9K-C9508-FM, and N9K-C9516-FM fabric modules support PIM6 ASM and SSM. Cisco Nexus 9500 Series switches with other line cards or fabric modules do not support PIM6.

- PIM bidirectional multicast source VLAN bridging is not supported on FEX ports.
- PIM6 Bidirectional is not supported.
- PIM6 is not supported on SVIs prior to Cisco NX-OS Release 9.3(3).
- PIM6 is not supported on any FEX ports (Layer 2 and Layer 3).
- PIM Bidirectional is supported for Cisco Nexus 9300-EX, Cisco Nexus 9300-FX/FX2/FX3 and Cisco Nexus 9300-GX platform switches.
- Cisco Nexus 9000 Series switches do not support PIM Bidir on vPCs or PIM6 ASM, SSM, and Bidirectional on vPCs.
- The following devices support PIM and PIM6 sparse mode on Layer 3 port-channel subinterfaces:
  - Cisco Nexus 9300 Series switches
  - Cisco Nexus 9300-EX Series switches and Cisco Nexus 3232C and 3264Q switches
  - Cisco Nexus 9500 Series switches with N9K-X9400 or N9K-X9500 line cards and/or N9K-C9504-FM, N9K-C9508-FM, and N9K-C9516-FM fabric modules.
- The multicast heavy template supports real-time packets and byte statistics but does not support VXLAN and tunnel egress or ingress statistics.
- Real-time/flex statistics is supported in:
  - Default template with configuration of **hardware profile multicast flex-stats-enable** command.
  - Heavy template without any configuration.

  Real-time statistics does not support ext-heavy template.
- GRE tunnels over IPv4 support multicast. GRE tunnels over IPv6 do not support multicast.
- Only Cisco Nexus 9300-EX and 9300-FX/FX2/FX3 platform switches support multicast on GRE tunnels.
- GRE tunnels does not support host connectivity.
- Because the IGMP functionality is not supported as part of the host connectivity, IGMP CLI is not available on GRE tunnels.
- You may not be able to add static tunnel OIFs to multicast routes, because IGMP CLI is not available on GRE tunnels, and it requires to statically bind a multicast group to the outgoing interface (OIF).
- Do not use SVI IP address as tunnel source or tunnel destination.
- Tunnel destination must be reachable via L3 physical interface or L3 subinterface.

- The L3 physical interface or subinterface via which the tunnel destination is reachable must be PIM enabled.
- Multiple GRE tunnels on the same device should not use the same source or the same destination.
- ECMP load sharing of GRE-encapsulated multicast traffic is not supported. If the tunnel destination is reachable across several links, the traffic is sent to only one of them.
- The multicast consistency checker is not supported on GRE tunnels.
- GRE tunnel can be a member of a VRF only if the source or destination interfaces are members of the same VRF.
- Multicast VRF-Lite Route Leaking is not supported for GRE.
- PIM Bidir is not supported with GRE.
- The Cisco Nexus 3232C and 3264Q switches do not support PIM6.
- When there is no PIM/PIM6 neighbor on an interface, the interface could be selected as an RPF interface based on the shortest/ECMP paths. Make sure to enable PIM/PIM6 on both the sides of the link when there are multiple ECMPs between the source and the receiver.
- Beginning with Cisco NX-OS Release 9.3(6), Multicast over GRE is supported on Cisco Nexus 9300-GX platform switches.
- Beginning with Cisco NX-OS Release 9.3(6), the following is supported:
  - Incoming RPF interface in Switch-1 is under default VRF and in Switch-2 on the other VRF.
  - Tunnel interface in Switch-1 is under default VRF and in Switch-2 on the other VRF.
  - Outgoing interface in Switch-1 is on the other VRF and in Switch-2 under default VRF.
- The presence of any GRE tunnel on the Cisco Nexus 9000 switches cannot co-exist with a sub-interface (multicast forwarding to a subinterface may be missing the dot1q tag). This impacts the receiving of multicast traffic on sub-interface. Traffic will be received at the parent interface and not at the sub-interface. This impact is only for regular/native multicast packets and not for Multicast GRE (encapsulation and decapsulation) packets. This limitation is applicable only to Cisco Nexus 9300-GX platform switches.
- The presence of any tunnel (feature tunnel or feature nv overlay) cannot co-exist with a sub-interface (multicast forwarding to a sub-interface may be missing the dot1q tag) this will impact receiving multicast traffic on a sub-interface. This limitation is applicable only to Cisco Nexus 9300-GX platform switches.
- In case GRE tunnel's sources or destinations were misconfigured (such as having incompatible sources/ destinations) they will be automatically shut down, and stay shut down even after the configuration has been recovered. The workaround is to manually shut/unshut such tunnels.
- In PIM-SM, some duplication or drops of packets are expected behavior when there are changes in the forwarding path. This behavior results in the following undesirable conditions:
  - When switching from receiving on the shared tree to shortest path tree (SPT), there is typically a small window when packets get dropped. The SPT feature may prevent this, but it may cause duplication sometimes.
  - The RP which initially forward packets that it may have received via PIM registers or MSDP will next join the SPT for native forwarding, and there is a small window where the RP may forward the same data packet twice, once as a native packet and once after PIM register or MSDP decap.

To resolve these issues, ensure that the forwarding path does not change by configuring a long (S,G) expiration time or by using SSM/PIM Bidir.

- PIM must be configured on all L3 interfaces between sources, receivers, and rendezvous points (RPs).
- HSRP-aware PIM is not supported in Cisco NX-OS.
- When a Multicast enabled BL is being reloaded, it is necessary to ensure that the customer has a Fabric port tracking timer set to around 3 - 5 depending on route scale.

  The fabric port tracking ensures that L3out comes up late and hence would delay the unicast convergence wrt the routes through the L3out. This would provide sufficient time for PIM overload timer (3 min) to complete and be ready to handle multicast states and data streams and Stripe winner related duties.

  Note that increasing port tracking would not affect any data streams currently running, since the expectation is that while the reloaded BL is still coming up, an alternate or backup BL is handling the data streams and performing Stripe winner related duties.

Guidelines and Limitations for Hello Messages

The following guidelines and limitations apply to Hello Messages:

- Default values for the PIM hello interval are recommended and should not be modified.

Guidelines and Limitations for Rendezvous Points

The following guidelines and limitations apply to Rendezvous Points (RP):

- Configure candidate RP intervals to a minimum of 15 seconds.
- Do not configure both Auto-RP and BSR protocols in the same network.
- PIM6 does not support BSRs and Auto-RP.
- You must configure PIM on the loopback interface that is used for the PIM Anycast RP and the PIM Bidir RP.
- The interface that is used to configure a PIM RP (whether static, BSR or Auto-RP) must have **ip** [**v6**] **pim sparse-mode**.
- To avoid excessive punts of the RPF failed packets, the Cisco Nexus 9000 Series switches may create S, G entries for active sources in ASM, although there is no rendezvous point (RP) for such group, or in situation when a reverse path forwarding (RPF) fails for the source.

  This behavior does not apply to Nexus 9200, 9300-EX platform switches, and N9K-X9700-EX LC platforms.
- If a device is configured with a BSR policy that should prevent it from being elected as the BSR, the device ignores the policy. This behavior results in the following undesirable conditions:
  - If a device receives a BSM that is permitted by the policy, the device, which incorrectly elected itself as the BSR, drops that BSM so that routers downstream fail to receive it. Downstream devices correctly filter the BSM from the incorrect BSR so that these devices do not receive RP information.
  - A BSM received by a BSR from a different device sends a new BSM but ensures that downstream devices do not receive the correct BSM.

- If the source VRF forwards multicast traffic across to a non-forwarder vPC peer which happens to be RP, then the S,G entries are not created on the forwarder vPC peer. This can lead to a drop in the multicast traffic for these sources. In order to avoid this, you must configure a anycast RP in the topology wherever the vPC peer is also a RP.

Guidelines and Limitations for Multicast VRF-lite Route Leaking

The following guidelines and limitations apply to multicast VRF-lite route leaking:

- Multicast VRF-lite route leaking is not supported on Cisco Nexus 9500 platform switches with -R or -RX line cards.

Default Settings

This table lists the default settings for PIM parameters.

Table 11: Default PIM Parameters

Parameters	Default
Use shared trees only	Disabled
Flush routes on restart	Disabled
Log neighbor changes	Disabled
Auto-RP message action	Disabled
BSR message action	Disabled
PIM sparse mode	Disabled
Designated router priority	1
Hello authentication mode	Disabled
Domain border	Disabled
RP address policy	No message filtering
PIM register message policy	No message filtering
BSR candidate RP policy	No message filtering
BSR policy	No message filtering
Auto-RP mapping agent policy	No message filtering
Auto-RP RP candidate policy	No message filtering
Join-prune policy	No message filtering

Parameters	Default
Neighbor adjacency policy	Become adjacent with all PIM neighbors
BFD	Disabled

Configuring PIM

Note Cisco NX-OS supports only PIM sparse mode version 2. In this publication, "PIM" refers to PIM sparse mode version 2.

You can configure separate ranges of addresses in the PIM domain using the multicast distribution modes described in the table below.

Multicast Distribution Mode	Requires RP Configuration	Description
ASM	Yes	Any source multicast
RPF routes for multicast	No	RPF routes for multicast

PIM Configuration Tasks

The following steps configure PIM .

1. Select the range of multicast groups that you want to configure in each multicast distribution mode.
2. Enable PIM .
3. Follow the configuration steps for the multicast distribution modes that you selected in Step 1.
   - For ASM mode, see Configuring ASM.
   - For RPF routes for multicast, see Configuring RPF Routes for Multicast.
4. Configure message filtering.

Note The CLI commands used to configure PIM are as follows:

- Configuration commands begin with **ip pim**.
- Show commands begin with **show ip pim**.

Enabling the PIM Feature

Before you can access the PIM commands, you must enable the PIM feature.

Before you begin

Ensure that you have installed the Enterprise Services license.

SUMMARY STEPS

1. **configure terminal**
2. **feature pim**
3. (Optional) **show running-configuration pim**
4. (Optional) **copy running-config startup-config**

DETAILED STEPS

Procedure

	Command or Action	Purpose

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

Step 2	**feature pim** **Example:** `switch(config)# feature pim`	Enables PIM. By default, PIM is disabled.
Step 3	(Optional) **show running-configuration pim** **Example:** `switch(config)# show running-configuration pim`	Shows the running-configuration information for PIM.
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 PIM Sparse Mode Parameters

You configure PIM sparse mode on every device interface that you want to participate in a sparse mode domain. You can configure the sparse mode parameters described in the table below.

Table 12: PIM Sparse Mode Parameters

Parameter	Description
Global to the device	

Parameter	Description
Auto-RP message action	Enables listening for and forwarding of Auto-RP messages. The default is disabled, which means that the router does not listen for or forward Auto-RP messages unless it is configured as a candidate RP or mapping agent.
BSR message action	Enables listening for and forwarding of BSR messages. The default is disabled, which means that the router does not listen for or forward BSR messages unless it is configured as a candidate RP or BSR candidate.
Register rate limit	Configures the IPv4 register rate limit in packets per second. The range is from 1 to 65,535. The default is no limit.
Initial holddown period	Configures the IPv4 initial holddown period in seconds. This holddown period is the time it takes for the MRIB to come up initially. If you want faster convergence, enter a lower value. The range is from 90 to 210. Specify 0 to disable the holddown period. The default is 210.
Per device interface	
PIM sparse mode	Enables PIM on an interface.
Designated router priority	Sets the designated router (DR) priority that is advertised in PIM hello messages on this interface. On a multi-access network with multiple PIM-enabled routers, the router with the highest DR priority is elected as the DR router. If the priorities match, the software elects the DR with the highest IP address. The DR originates PIM register messages for the directly connected multicast sources and sends PIM join messages toward the rendezvous point (RP) for directly connected receivers. Values range from 1 to 4294967295. The default is 1.
Designated router delay	Delays participation in the designated router (DR) election by setting the DR priority that is advertised in PIM hello messages to 0 for a specified period. During this delay, no DR changes occur, and the current switch is given time to learn all of the multicast states on that interface. After the delay period expires, the correct DR priority is sent in the hello packets, which retriggers the DR election. Values range from 3 to 0xffff seconds.
Hello authentication mode	Enables an MD5 hash authentication key, or password, in PIM hello messages on the interface so that directly connected neighbors can authenticate each other. The PIM hello messages are IPsec encoded using the Authentication Header (AH) option. You can enter an unencrypted (cleartext) key or one of these values followed by a space and the MD5 authentication key: • 0—Specifies an unencrypted (cleartext) key • 3—Specifies a 3-DES encrypted key • 7—Specifies a Cisco Type 7 encrypted key The authentication key can be up to 16 characters. The default is disabled.

Parameter	Description
Hello interval	Configures the interval at which hello messages are sent in milliseconds. The range is from 1000 to 18724286. The default is 30000. **Note** See the *Cisco Nexus 9000 Series NX-OS Verified Scalability Guide* for the verified range of this parameter and associated PIM neighbor scale.
Domain border	Enables the interface to be on the border of a PIM domain so that no bootstrap, candidate-RP, or Auto-RP messages are sent or received on the interface. The default is disabled.
Neighbor policy	Configures which PIM neighbors to become adjacent to based on a prefix-list policy.[3] If the policy name does not exist or no prefix lists are configured in a policy, adjacency is established with all neighbors. The default is to become adjacent with all PIM neighbors. **Note** We recommend that you should configure this feature only if you are an experienced network administrator. **Note** The PIM neighbor policy supports only prefix lists. It does not support ACLs used inside a route map.

[3] To configure prefix-list policies, see the *Cisco Nexus 9000 Series NX-OS Unicast Routing Configuration Guide*.

Configuring PIM Sparse Mode Parameters

SUMMARY STEPS

1. **configure terminal**
2. (Optional) **ip pim auto-rp** {**listen** [**forward**] | **forward** [**listen**]}
3. (Optional) **ip pim bsr** {**listen** [**forward**] | **forward** [**listen**]}
4. (Optional) **ip pim register-rate-limit** *rate*
5. (Optional) **ip pim spt-threshold infinity group-list** *route-map-name*
6. (Optional) [**ip** | **ipv4**] **routing multicast holddown** *holddown-period*
7. (Optional) **show running-configuration pim**
8. **interface** *interface*
9. **ip pim sparse-mode**
10. (Optional) **ip pim dr-priority** *priority*
11. (Optional) **ip pim dr-delay** *delay*
12. (Optional) **ip pim hello-authentication ah-md5** *auth-key*
13. (Optional) **ip pim hello-interval** *interval*
14. (Optional) **ip pim border**
15. (Optional) **ip pim neighbor-policy prefix-list** *prefix-list*
16. (Optional) **show ip pim interface** [*interface* | **brief**] [**vrf** *vrf-name* | **all**]

17. (Optional) **copy running-config startup-config**

DETAILED STEPS

Procedure

	Command or Action	Purpose

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

Step 2	(Optional) **ip pim auto-rp** {**listen** [**forward**] | **forward** [**listen**]} **Example:** `switch(config)# ip pim auto-rp listen`	Enables listening for or forwarding of Auto-RP messages. The default is disabled, which means that the software does not listen for or forward Auto-RP messages.
Step 3	(Optional) **ip pim bsr** {**listen** [**forward**] | **forward** [**listen**]} **Example:** `switch(config)# ip pim bsr forward`	Enables listening for or forwarding of BSR messages. The default is disabled, which means that the software does not listen for or forward BSR messages.
Step 4	(Optional) **ip pim register-rate-limit** *rate* **Example:** `switch(config)# ip pim register-rate-limit 1000`	Configures the rate limit in packets per second. The range is from 1 to 65,535. The default is no limit.
Step 5	(Optional) **ip pim spt-threshold infinity group-list** *route-map-name* **Example:** `switch(config)# ip pim spt-threshold infinity group-list my_route-map-name`	Creates the IPv4 PIM (*, G) state only, for the group prefixes defined in the specified route map. Cisco NX-OS Release 3.1 supports up to 1000 route-map entries, and Cisco NX-OS releases prior to 3.1 support up to 500 route-map entries. **Note** The **ip pim use-shared-tree-only group-list** command performs the same function as the **ip pim spt-threshold infinity group-list** command. You can choose to use either command to implement this step. Both the commands (**ip pim spt-threshold infinity group-list** and **ip pim use-shared-tree-only group-list** has the following limitations: • It is only supported for virtual port channels (vPC) on the Cisco Nexus 9000 Cloud Scale Switches. • It is supported in NX-OS (non-vPC) Last Hop Router (LHR) configurations.

	Command or Action	Purpose
Step 6	(Optional) [**ip** | **ipv4**] **routing multicast holddown** *holddown-period* **Example:** `switch(config)# ip routing multicast holddown 100`	Configures the initial holddown period in seconds. The range is from 90 to 210. Specify 0 to disable the holddown period. The default is 210.
Step 7	(Optional) **show running-configuration pim** **Example:** `switch(config)# show running-configuration pim`	Displays PIM running-configuration information.

Step 8	**interface** *interface*
	Example:
	`switch(config)# interface ethernet 2/1`
	`switch(config-if)#`
		Enters interface configuration mode.

Step 9	**ip pim sparse-mode** **Example:** `switch(config-if)# ip pim sparse-mode`	Enables PIM sparse mode on this interface. The default is disabled.
Step 10	(Optional) **ip pim dr-priority** *priority* **Example:** `switch(config-if)# ip pim dr-priority 192`	Sets the designated router (DR) priority that is advertised in PIM hello messages. Values range from 1 to 4294967295. The default is 1.
Step 11	(Optional) **ip pim dr-delay** *delay* **Example:** `switch(config-if)# ip pim dr-delay 3`	Delays participation in the designated router (DR) election by setting the DR priority that is advertised in PIM hello messages to 0 for a specified period. During this delay, no DR changes occur, and the current switch is given time to learn all of the multicast states on that interface. After the delay period expires, the correct DR priority is sent in the hello packets, which retriggers the DR election. Values range from 3 to 0xffff seconds. **Note** This command delays participation in the DR election only upon bootup or following an IP address or interface state change. It is intended for use with multicast-access non-vPC Layer 3 interfaces only.
Step 12	(Optional) **ip pim hello-authentication ah-md5** *auth-key* **Example:** `switch(config-if)# ip pim hello-authentication ah-md5 my_key`	Enables an MD5 hash authentication key in PIM hello messages. You can enter an unencrypted (cleartext) key or one of these values followed by a space and the MD5 authentication key: • 0—Specifies an unencrypted (cleartext) key • 3—Specifies a 3-DES encrypted key • 7—Specifies a Cisco Type 7 encrypted key The key can be up to 16 characters. The default is disabled.

	Command or Action	Purpose
Step 13	(Optional) **ip pim hello-interval** *interval* **Example:** `switch(config-if)# ip pim hello-interval 25000`	Configures the interval at which hello messages are sent in milliseconds. The range is from 1000 to 18724286. The default is 30000. **Note** The minimum value is 1 millisecond.
Step 14	(Optional) **ip pim border** **Example:** `switch(config-if)# ip pim border`	Enables the interface to be on the border of a PIM domain so that no bootstrap, candidate-RP, or Auto-RP messages are sent or received on the interface. The default is disabled.
Step 15	(Optional) **ip pim neighbor-policy prefix-list** *prefix-list* **Example:** `switch(config-if)# ip pim neighbor-policy prefix-list AllowPrefix`	Enables the interface to be on the border of a PIM domain so that no bootstrap, candidate-RP, or Auto-RP messages are sent or received on the interface. The default is disabled. Also configures which PIM neighbors to become adjacent to based on a prefix-list policy with the **ip prefix-list** *prefix-list* command. The prefix list can be up to 63 characters. The default is to become adjacent with all PIM neighbors. **Note** We recommend that you configure this feature only if you are an experienced network administrator.
Step 16	(Optional) **show ip pim interface** [*interface* | **brief**] [**vrf** *vrf-name* | **all**] **Example:** `switch(config-if)# show ip pim interface`	Displays PIM interface information.
Step 17	(Optional) **copy running-config startup-config** **Example:** `switch(config-if)# copy running-config startup-config`	Copies the running configuration to the startup configuration.

Configuring PIM6 Sparse Mode Parameters

SUMMARY STEPS

1. **configure terminal**
2. (Optional) **ipv6 pim register-rate-limit** *rate*
3. (Optional) **ipv6 routing multicast holddown** *holddown-period*
4. (Optional) **show running-configuration pim6**
5. **interface** *interface*
6. **ipv6 pim sparse-mode**
7. (Optional) **ipv6 pim dr-priority** *priority*

8. (Optional) **ipv6 pim hello-interval** *interval*
9. (Optional) **ipv6 pim border**
10. (Optional) **ipv6 pim neighbor-policy prefix-list** *prefix-list*
11. **show ipv6 pim interface** [*interface* | brief] [vrf *vrf-name* | all]
12. **copy running-config startup-config**

DETAILED STEPS

Procedure

	Command or Action	Purpose

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

Step 2	(Optional) **ipv6 pim register-rate-limit** *rate* **Example:** `switch(config)# ipv6 pim register-rate-limit 1000`	Configures the rate limit in packets per second. The range is from 1 to 65,535. The default is no limit.
Step 3	(Optional) **ipv6 routing multicast holddown** *holddown-period* **Example:** `switch(config)# ipv6 routing multicast holddown 100`	Configures the initial holddown period in seconds. The range is from 90 to 210. Specify 0 to disable the holddown period. The default is 210.
Step 4	(Optional) **show running-configuration pim6** **Example:** `switch(config)# show running-configuration pim6`	Displays PIM6 running-configuration information, including the register rate limit.

Step 5	**interface** *interface*
	Example:
	`switch(config)# interface vlan 10`
	`switch(config-if)#`
		Enters interface configuration mode on the specified interface.

Step 6	**ipv6 pim sparse-mode** **Example:** `switch(config-if)# ipv6 pim sparse-mode`	Enables PIM sparse mode on this interface. The default is disabled. Beginning with Cisco NX-OS Release 9.3(5) you can configure this command on a SVI interface in Broadcom-based switches.
Step 7	(Optional) **ipv6 pim dr-priority** *priority* **Example:** `switch(config-if)# ipv6 pim dr-priority 192`	Sets the designated router (DR) priority that is advertised in PIM6 hello messages. Values range from 1 to 4294967295. The default is 1.

	Command or Action	Purpose
Step 8	(Optional) **ipv6 pim hello-interval** *interval* **Example:** `switch(config-if)# ipv6 pim hello-interval 25000`	Configures the interval at which hello messages are sent in milliseconds. The range is from 1000 to 18724286. The default is 30000.
Step 9	(Optional) **ipv6 pim border** **Example:** `switch(config-if)# ipv6 pim border`	Enables the interface to be on the border of a PIM6 domain so that no bootstrap, candidate-RP, or Auto-RP messages are sent or received on the interface. The default is disabled.
Step 10	(Optional) **ipv6 pim neighbor-policy prefix-list** *prefix-list* **Example:** `switch(config-if)# ipv6 pim neighbor-policy prefix-list AllowPrefix`	Configures which PIM6 neighbors to become adjacent to based on a prefix-list policy with the **ipv6 prefix-list** *prefix-list* command. The prefix list can be up to 63 characters. The default is to become adjacent with all PIM6 neighbors. **Note** We recommend that you configure this feature only if you are an experienced network administrator.
Step 11	**show ipv6 pim interface** [*interface* | brief] [vrf *vrf-name* | all] **Example:** `switch(config-if)# show ipv6 pim interface`	Displays PIM6 interface information.
Step 12	**copy running-config startup-config** **Example:** `switch(config-if)# copy running-config startup-config`	(Optional) Saves configuration changes.

Configuring ASM

To configure ASM mode, you configure sparse mode and the RP selection method, where you indicate the distribution mode and assign the range of multicast groups.

Configuring Static RPs

You can configure an RP statically by configuring the RP address on every router that will participate in the PIM domain.

Note We recommend that the RP address uses the loopback interface and also the interface with the RP address must have **ip pim sparse-mode** enabled.

You can specify a route-map policy name that lists the group prefixes to use with the **match ip multicast** command or specify a prefix-list method of configuration.

Note Cisco NX-OS always uses the longest-match prefix to find the RP, so the behavior is the same irrespective of the position of the group prefix in the route map or in the prefix list.

The following example configuration produces the same output using Cisco NX-OS (231.1.1.0/24 is always denied irrespective of the sequence number):

```
ip prefix-list plist seq 10 deny 231.1.1.0/24
ip prefix-list plist seq 20 permit 231.1.0.0/16
ip prefix-list plist seq 10 permit 231.1.0.0/16
ip prefix-list plist seq 20 deny 231.1.1.0/24
```

Configuring Static RPs

Before you begin

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

SUMMARY STEPS

1. **configure terminal**
2. **ip pim rp-address** *rp-address* [**group-list** *ip-prefix* | **prefix-list** *name* |**override** |**route-map** *policy-name*] [**bidir**]
3. (Optional) **show ip pim group-range** [*ip-prefix* | **vrf** *vrf-name*]
4. (Optional) **copy running-config startup-config**

DETAILED STEPS

Procedure

	Command or Action	Purpose

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

Step 2	**ip pim rp-address** *rp-address* [**group-list** *ip-prefix*	**prefix-list** *name* |**override** |**route-map** *policy-name*] [**bidir**] **Example:** `switch(config)# ip pim rp-address 192.0.2.33 group-list 224.0.0.0/9`	Configures a PIM static RP address for a multicast group range. You can specify a prefix-list policy name for the static RP address or a route-map policy name that lists the group prefixes to use with the **match ip multicast** command. The mode is ASM. The **override** option causes the RP address to override the dynamically learned RP addresses for specified groups in route-map. The example configures PIM ASM mode for the specified group range.

	Command or Action	Purpose
Step 3	(Optional) **show ip pim group-range** [*ip-prefix* | **vrf** *vrf-name*] **Example:** `switch(config)# show ip pim group-range`	Displays PIM RP information, including BSR listen and forward states.
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 Static RPs (PIM6)

Before you begin

Ensure that you have installed the Enterprise Services license and enabled PIM6.

SUMMARY STEPS

1. **configure terminal**
2. **ipv6 pim rp-address** *rp-address* [**group-list** *ipv6-prefix* | **route-map** *policy-nsmr*]
3. (Optional) **show ipv6 pim group-range** [*ipv6-prefix* | **vrf** *vrf-name*]
4. (Optional) **copy running-config startup-config**

DETAILED STEPS

Procedure

	Command or Action	Purpose

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

Step 2	**ipv6 pim rp-address** *rp-address* [**group-list** *ipv6-prefix* | **route-map** *policy-nsmr*] **Example:** `switch(config)# ipv6 pim rp-address 2001:0db8:0:abcd::1 group-list ff1e:abcd:def1::0/24`	Configures a PIM6 static RP address for a multicast group range. You can specify a route-map policy name that lists the group prefixes to use with the **match ip multicast** command. The mode is ASM. The default group range is ff00::0/8. The example configures PIM6 ASM mode for the specified group range.
Step 3	(Optional) **show ipv6 pim group-range** [*ipv6-prefix* | **vrf** *vrf-name*] **Example:** `switch(config)# show ipv6 pim group-range`	Displays PIM6 modes and group ranges.

	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.

Configuring BSRs

You configure BSRs by selecting candidate BSRs and RPs.

Caution Do not configure both Auto-RP and BSR protocols in the same network.

You can configure a candidate BSR with the arguments described in the table below.

Table 13: Candidate BSR Arguments

Argument	Description
interface	Interface type and number used to derive the BSR source IP address used in bootstrap messages.
hash-length	Number of high order 1s used to form a mask that is ANDed with group address ranges of candidate RPs to form a hash value. The mask determines the number of consecutive addresses to assign across RPs with the same group range. For PIM, this value ranges from 0 to 32 and has a default of 30.
priority	Priority assigned to this BSR. The software elects the BSR with the highest priority, or if the BSR priorities match, the software elects the BSR with the highest IP address. This value ranges from 0, the lowest priority, to 255 and has a default of 64.

Configuring BSRs Candidate RP Arguments and Keywords

You can configure a candidate RP with the arguments and keywords described in this table.

Table 14: BSR Candidate RP Arguments and Keywords

Argument or Keyword	Description
interface	Interface type and number used to derive the BSR source IP address used in bootstrap messages.
group-list *ip-prefix*	Multicast groups handled by this RP specified in a prefix format.
interval	Number of seconds between sending candidate-RP messages. This value ranges from 1 to 65,535 and has a default of 60 seconds. **Note** We recommend that you configure the candidate RP interval to a minimum of 15 seconds.

Argument or Keyword	Description
priority	Priority assigned to this RP. The software elects the RP with the highest priority a range of groups or, if the priorities match, the highest IP address. (The highest priority is the lowest numerical value.) This value ranges from 0, the highest prio to 255 and has a default of 192. **Note** This priority differs from the BSR BSR-candidate priority, which prefers the high value between 0 and 255.
route-map *policy-name*	Route-map policy name that defines the group prefixes where this feature is app

Tip You should choose the candidate BSRs and candidate RPs that have good connectivity to all parts of the PIM domain.

You can configure the same router to be both a BSR and a candidate RP. In a domain with many routers, you can select multiple candidate BSRs and RPs to automatically fail over to alternates if a BSR or an RP fails.

To configure candidate BSRs and RPs, follow these steps:

1. Configure whether each router in the PIM domain should listen for and forward BSR messages. A router configured as either a candidate RP or a candidate BSR will automatically listen for and forward all bootstrap router protocol messages, unless an interface is configured with the domain border feature.
2. Select the routers to act as candidate BSRs and RPs.
3. Configure each candidate BSR and candidate RP as described in this section.
4. Configure BSR message filtering.

Configuring BSRs

Before you begin

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

SUMMARY STEPS

1. **configure terminal**
2. **ip pim bsr** {**forward** [**listen**] | **listen** [**forward**]}
3. **ip pim bsr bsr-candidate** *interface* [**hash-len** *hash-length*] [**priority** *priority*]
4. **ip pim sparse-mode**
5. (Optional) **ip pim bsr rp-candidate** *interface* **group-list** *ip-prefix* **route-map** *policy-name* **priority** *priority* **interval** *interval*
6. (Optional) **show ip pim group-range** [*ip-prefix* | **vrf** *vrf-name*]
7. (Optional) **copy running-config startup-config**

DETAILED STEPS

Procedure

	Command or Action	Purpose

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

Step 2	**ip pim bsr** {**forward** [**listen**] | **listen** [**forward**]} **Example:** `switch(config)# ip pim bsr listen forward`	Configures listen and forward. Ensure that you have entered this command in each VRF on the remote PE.
Step 3	**ip pim bsr bsr-candidate** *interface* [**hash-len** *hash-length*] [**priority** *priority*] **Example:** `switch(config)# ip pim bsr-candidate ethernet 2/1 hash-len 24`	Configures a candidate bootstrap router (BSR). The source IP address used in a bootstrap message is the IP address of the interface. The hash length ranges from 0 to 32 and has a default of 30. The priority ranges from 0 to 255 and has a default of 64.
Step 4	**ip pim sparse-mode** **Example:** `switch(config-if)# ip pim sparse-mode`	Enables PIM sparse mode on this interface. The default is disabled.
Step 5	(Optional) **ip pim bsr rp-candidate** *interface* **group-list** *ip-prefix* **route-map** *policy-name* **priority** *priority* **interval** *interval* **Example:** `switch(config)# ip pim rp-candidate ethernet 2/1 group-list 239.0.0.0/24`	Configures a candidate RP for BSR. The priority ranges from 0, the highest priority, to 65,535 and has a default of 192. The interval ranges from 1 to 65,535 seconds and has a default of 60. **Note** We recommend that you configure the candidate RP interval to a minimum of 15 seconds. The example configures an ASM candidate RP.
Step 6	(Optional) **show ip pim group-range** [*ip-prefix* | **vrf** *vrf-name*] **Example:** `switch(config)# show ip pim group-range`	Displays PIM modes and group ranges.
Step 7	(Optional) **copy running-config startup-config** **Example:** `switch(config)# copy running-config startup-config`	Copies the running configuration to the startup configuration.

Configuring Auto-RP

You can configure Auto-RP by selecting candidate mapping agents and RPs. You can configure the same router to be both a mapping agent and a candidate RP.

Caution Do not configure both Auto-RP and BSR protocols in the same network.

You can configure an Auto-RP mapping agent with the arguments described in this table.

Table 15: Auto-RP Mapping Agent Arguments

Argument	Description
interface	Interface type and number used to derive the IP address of the Auto-RP mapping agent used in bootstrap messages.
scope *ttl*	Time-to-Live (TTL) value that represents the maximum number of hops that RP-Discovery messages are forwarded. This value can range from 1 to 255 and has a default of 32.

If you configure multiple Auto-RP mapping agents, only one is elected as the mapping agent for the domain. The elected mapping agent ensures that all candidate RP messages are sent out. All mapping agents receive the candidate RP messages and advertise the same RP cache in their RP-discovery messages.

You can configure a candidate RP with the arguments and keywords described in this table.

Table 16: Auto-RP Candidate RP Arguments and Keywords

Argument or Keyword	Description
interface	Interface type and number used to derive the IP address of the candidate RP used in bootstrap messages.
group-list *ip-prefix*	Multicast groups handled by this RP. It is specified in a prefix format.
scope *ttl*	Time-to-Live (TTL) value that represents the maximum number of hops that RP-Discovery messages are forwarded. This value can range from 1 to 255 and has a default of 32.
interval	Number of seconds between sending RP-Announce messages. This valu can range from 1 to 65,535 and has a default of 60. **Note** We recommend that you configure the candidate RP interval to a minimur of 15 seconds.
route-map *policy-name*	Route-map policy name that defines the group prefixes where this featur is applied.

Tip You should choose mapping agents and candidate RPs that have good connectivity to all parts of the PIM domain.

To configure Auto-RP mapping agents and candidate RPs, follow these steps:

1. For each router in the PIM domain, configure whether that router should listen for and forward Auto-RP messages. A router configured as either a candidate RP or an Auto-RP mapping agent will automatically

listen for and forward all Auto-RP protocol messages, unless an interface is configured with the domain border feature.

2. Select the routers to act as mapping agents and candidate RPs.
3. Configure each mapping agent and candidate RP as described in this section.
4. Configure Auto-RP message filtering.

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

Configuring Auto RP (PIM)

Before you begin

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

SUMMARY STEPS

1. **configure terminal**
2. **ip pim** {**send-rp-discovery** | **auto-rp mapping-agent**} *interface* [**scope** *ttl*]
3. **ip pim** {**send-rp-announce** | **auto-rp rp-candidate**} *interface* {**group-list** *ip-prefix* | **prefix-list** *name* | **route-map** *policy-name*} [**scope** *ttl*] **interval** *interval*] [**bidir**]
4. **ip pim sparse-mode**
5. (Optional) **show ip pim group-range** [*ip-prefix* | **vrf** *vrf-name*]
6. (Optional) **copy running-config startup-config**

DETAILED STEPS

Procedure

	Command or Action	Purpose

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

Step 2	**ip pim** {**send-rp-discovery**	**auto-rp mapping-agent**} *interface* [**scope** *ttl*] **Example:** `switch(config)# ip pim auto-rp mapping-agent ethernet 2/1`	Configures an Auto-RP mapping agent. The source IP address used in Auto-RP Discovery messages is the IP address of the interface. The default scope is 32.
Step 3	**ip pim** {**send-rp-announce**	**auto-rp rp-candidate**} *interface* {**group-list** *ip-prefix*	**prefix-list** *name*	**route-map** *policy-name*} [**scope** *ttl*] **interval** *interval*] [**bidir**] **Example:** `switch(config)# ip pim auto-rp rp-candidate ethernet 2/1 group-list 239.0.0.0/24`	Configures an Auto-RP candidate RP. The default scope is 32. The default interval is 60 seconds. By default, the command creates an ASM candidate RP. Use the **bidir** option to create a Bidir candidate RP. **Note** We recommend that you configure the candidate RP interval to a minimum of 15 seconds.

	Command or Action	Purpose
		The example configures an ASM candidate RP.
Step 4	**ip pim sparse-mode** **Example:** `switch(config-if)# ip pim sparse-mode`	Enables PIM sparse mode on this interface. The default is disabled.
Step 5	(Optional) **show ip pim group-range** [*ip-prefix* | **vrf** *vrf-name*] **Example:** `switch(config)# show ip pim group-range`	Displays PIM modes and group ranges.
Step 6	(Optional) **copy running-config startup-config** **Example:** `switch(config)# copy running-config startup-config`	Copies the running configuration to the startup configuration.

Configuring a PIM Anycast-RP Set

To configure a PIM Anycast-RP set, follow these steps:

1. Select the routers in the PIM Anycast-RP set.
2. Select an IP address for the PIM Anycast-RP set.
3. Configure each peer RP in the PIM Anycast-RP set as described in this section.

Configuring a PIM Anycast RP Set

Before you begin

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

SUMMARY STEPS

1. **configure terminal**
2. **interface loopback** *number*
3. **ip address** *ip-prefix*
4. **ip pim sparse-mode**
5. **ip router** *routing-protocol-configuration*
6. **exit**
7. **interface loopback** *number*
8. **ip address** *ip-prefix*
9. **ip pim sparse-mode**
10. **ip router** *routing-protocol-configuration*
11. **exit**
12. **ip pim rp-address** *anycast-rp-address* [**group-list** *ip-address*]
13. **ip pim anycast-rp** *anycast-rp-address anycast-rp-set-router-address*

14. Repeat Step 13 using the same Anycast-RP address for each peer router in the RP set (including the local router).
15. (Optional) **show ip pim rp**
16. (Optional) **show ip mroute** *ip-address*
17. (Optional) **show ip pim group-range** [*ip-prefix* | **vrf** *vrf-name*]
18. (Optional) **copy running-config startup-config**

DETAILED STEPS

Procedure

	Command or Action	Purpose

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

Step 2	**interface loopback** *number*
	Example:
	`switch(config)# interface loopback 0`
	`switch(config-if)#`
		Configures an interface loopback.
		This example configures interface loopback 0.

Step 3	**ip address** *ip-prefix* **Example:** `switch(config-if)# ip address 192.168.1.1/32`	Configures an IP address for this interface. It should be a unique IP address that helps to identify this router.
Step 4	**ip pim sparse-mode** **Example:** `switch(config-if)# ip pim sparse-mode`	Enables PIM sparse mode.
Step 5	**ip router** *routing-protocol-configuration* **Example:** `switch(config-if)# ip router ospf 1 area 0.0.0.0`	Enables the interface to be reachable by other routers in the Anycast RP set.

Step 6	**exit**
	Example:
	`switch(config-if)# exit`
	`switch(config)#`
		Exits interface configuration mode.

Step 7	**interface loopback** *number*
	Example:
	`switch(config)# interface loopback 1`
	`switch(config-if)#`
		Configures an interface loopback.
		This example configures interface loopback 1.

Step 8	**ip address** *ip-prefix* **Example:**	Configures an IP address for this interface. It should be a common IP address that acts as the Anycast RP address.

	Command or Action	Purpose
	`switch(config-if)# ip address 10.1.1.1/32`	
Step 9	**ip pim sparse-mode** **Example:** `switch(config-if)# ip pim sparse-mode`	Enables PIM sparse mode on this interface. The default is disabled.
Step 10	**ip router** *routing-protocol-configuration* **Example:** `switch(config-if)# ip router ospf 1 area 0.0.0.0`	Enables the interface to be reachable by other routers in the Anycast RP set.

Step 11	**exit**
	Example:
	`switch(config-if)# exit`
	`switch(config)#`
		Exits interface configuration mode.

Step 12	**ip pim rp-address** *anycast-rp-address* [**group-list** *ip-address*] **Example:** `switch(config)# ip pim rp-address 10.1.1.1 group-list 224.0.0.0/4`	Configures the PIM Anycast RP address.
Step 13	**ip pim anycast-rp** *anycast-rp-address anycast-rp-set-router-address* **Example:** `switch(config)# ip pim anycast-rp 10.1.1.1 192.168.1.1`	Configures a PIM Anycast-RP peer address for the specified Anycast-RP address. Each command with the same Anycast-RP address forms an Anycast-RP set. The IP addresses of RPs are used for communication with RPs in the set.
Step 14	Repeat Step 13 using the same Anycast-RP address for each peer router in the RP set (including the local router).	—
Step 15	(Optional) **show ip pim rp** **Example:** `switch(config)# show ip pim rp`	Displays the PIM RP mapping.
Step 16	(Optional) **show ip mroute** *ip-address* **Example:** `switch(config)# show ip mroute 239.1.1.1`	Displays the mroute entries.
Step 17	(Optional) **show ip pim group-range** [*ip-prefix* | **vrf** *vrf-name*] **Example:** `switch(config)# show ip pim group-range`	Displays PIM modes and group ranges.
Step 18	(Optional) **copy running-config startup-config** **Example:** `switch(config)# copy running-config startup-config`	Copies the running configuration to the startup configuration.

Configuring a PIM Anycast RP Set (PIM6)

Before you begin

Ensure that you have installed the Enterprise Services license and enabled PIM6.

SUMMARY STEPS

1. **configure terminal**
2. **interface loopback** *number*
3. **ipv6 address** *ipv6-prefix*
4. **ipv6 pim sparse-mode**
5. **ipv6 router** *routing-protocol-configuration*
6. **exit**
7. **interface loopback** *number*
8. **ipv6 address** *ipv6-prefix*
9. **ipv6 router** *routing-protocol-configuration*
10. **exit**
11. **ipv6 pim rp-address** *anycast-rp-address* [**group-list** *ip-address*]
12. **ipv6 pim anycast-rp** *anycast-rp-address anycast-rp-set-router-address*
13. Repeat Step 13 using the same Anycast-RP address for each peer router in the RP set (including the local router).
14. (Optional) **show ipv6 pim rp**
15. (Optional) **show ipv6 mroute** *ipv6-address*
16. (Optional) **show ipv6 pim group-range** [*ipv6-prefix*] [**vrf** *vrf-name* | **all**]
17. (Optional) **copy running-config startup-config**

DETAILED STEPS

Procedure

	Command or Action	Purpose

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

Step 2	**interface loopback** *number*
	Example:
	`switch(config)# interface loopback 0`
	`switch(config-if)#`
		Configures an interface loopback.
		This example configures interface loopback 0.

Step 3	**ipv6 address** *ipv6-prefix* **Example:** `switch(config-if)# ipv6 address 2001:0db8:0:abcd::5/32`	Configures an IP address for this interface. It should be a unique IP address that helps to identify this router.

	Command or Action	Purpose
Step 4	**ipv6 pim sparse-mode** **Example:** `switch(config-if)# ipv6 pim sparse-mode`	Enable PIM6 sparse mode.
Step 5	**ipv6 router** *routing-protocol-configuration* **Example:** `switch(config-if)# ipv6 router ospfv3 1 area 0.0.0.0`	Enables the interface to be reachable by other routers in the Anycast RP set.

Step 6	**exit**
	Example:
	`switch(config-if)# exit`
	`switch(config)#`
		Exits interface configuration mode.

Step 7	**interface loopback** *number*
	Example:
	`switch(config)# interface loopback 1`
	`switch(config-if)#`
		Configures an interface loopback.
		This example configures interface loopback 1.

Step 8	**ipv6 address** *ipv6-prefix* **Example:** `switch(config-if)# ipv6 address 2001:0db8:0:abcd::1111/32`	Configures an IP address for this interface. It should be a common IP address that acts as the Anycast RP address.
Step 9	**ipv6 router** *routing-protocol-configuration* **Example:** `switch(config-if)# ipv6 router ospfv3 1 area 0.0.0.0`	Enables the interface to be reachable by other routers in the Anycast RP set.

Step 10	**exit**
	Example:
	`switch(config-if)# exit`
	`switch(config)#`
		Exits interface configuration mode.

Step 11	**ipv6 pim rp-address** *anycast-rp-address* [**group-list** *ip-address*] **Example:** `switch(config)# ipv6 pim rp-address 2001:0db8:0:abcd::1111 group-list ff1e:abcd:def1::0/24`	Configures the PIM6 Anycast RP address.
Step 12	**ipv6 pim anycast-rp** *anycast-rp-address anycast-rp-set-router-address* **Example:** `switch(config)# ipv6 pim anycast-rp 2001:0db8:0:abcd::5 2001:0db8:0:abcd::1111`	Configures a PIM6 Anycast-RP peer address for the specified Anycast-RP address. Each command with the same Anycast-RP address forms an Anycast-RP set. The IP addresses of RPs are used for communication with RPs in the set.

	Command or Action	Purpose
Step 13	Repeat Step 13 using the same Anycast-RP address for each peer router in the RP set (including the local router).	—
Step 14	(Optional) **show ipv6 pim rp** **Example:** `switch(config)# show ipv6 pim rp`	Displays the PIM RP mapping.
Step 15	(Optional) **show ipv6 mroute** *ipv6-address* **Example:** `switch(config)# show ipv6 mroute ff1e:2222::1:1:1:1`	Displays the mroute entries.
Step 16	(Optional) **show ipv6 pim group-range** [*ipv6-prefix*] [**vrf** *vrf-name* | **all**] **Example:** `switch(config)# show ipv6 pim group-range`	Displays PIM6 modes and group ranges.
Step 17	(Optional) **copy running-config startup-config** **Example:** `switch(config)# copy running-config startup-config`	Copies the running configuration to the startup configuration.

Configuring Shared Trees Only for ASM

You can configure shared trees only on the last-hop router for Any Source Multicast (ASM) groups, which means that the router never switches over from the shared tree to the SPT when a receiver joins an active group. You can specify a group range where the use of shared trees is to be enforced with the **match ip multicast** command. This option does not affect the normal operation of the router when a source tree join-prune message is received.

Note The Cisco NX-OS software does not support the shared-tree feature on vPCs. For more information about vPCs, see the *Cisco Nexus 9000 Series NX-OS Interfaces Configuration Guide*.

The default is disabled, which means that the software can switch over to source trees.

Note In ASM mode, only the last-hop router switches from the shared tree to the SPT.

Configuring Shared Trees Only for ASM

Before you begin

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

SUMMARY STEPS

1. **configure terminal**
2. **ip pim use-shared-tree-only group-list** *policy-name*
3. (Optional) **show ip pim group-range** [*ip-prefix* | **vrf** *vrf-name*]
4. (Optional) **copy running-config startup-config**

DETAILED STEPS

Procedure

	Command or Action	Purpose

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

Step 2	**ip pim use-shared-tree-only group-list** *policy-name* **Example:** `switch(config)# ip pim use-shared-tree-only group-list my_group_policy`	Builds only shared trees, which means that the software never switches over from the shared tree to the SPT. You specify a route-map policy name that lists the groups to use with the **match ip multicast** command. By default, the software triggers a PIM (S, G) join toward the source when it receives multicast packets for a source for which it has the (*, G) state. This command has the following limitations: • It is only supported for virtual port channels (vPC) on the Cisco Nexus 9000 Cloud Scale Switches. • It is supported in NX-OS (non-vPC) Last Hop Router (LHR) configurations.
Step 3	(Optional) **show ip pim group-range** [*ip-prefix* | **vrf** *vrf-name*] **Example:** `switch(config)# show ip pim group-range`	Displays PIM modes and group ranges.
Step 4	(Optional) **copy running-config startup-config** **Example:** `switch(config-if)# copy running-config startup-config`	Copies the running configuration to the startup configuration.

Configuring Shared Trees Only for ASM (PIM6)

Before you begin

Ensure that you have installed the Enterprise Services license and enabled PIM6.

SUMMARY STEPS

1. **configure terminal**
2. **ipv6 pim use-shared-tree-only group-list** *policy-name*
3. (Optional) **show ipv6 pim group-range** [*ipv6-prefix* | **vrf** *vrf-name*]
4. (Optional) **copy running-config startup-config**

DETAILED STEPS

Procedure

	Command or Action	Purpose

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

Step 2	**ipv6 pim use-shared-tree-only group-list** *policy-name* **Example:** `switch(config)# ipv6 pim use-shared-tree-only group-list my_group_policy`	Builds only shared trees, which means that the software never switches over from the shared tree to the SPT. You specify a route-map policy name that lists the groups to use with the **match ipv6 multicast** command. By default, the software triggers a PIM (S, G) join toward the source when it receives multicast packets for a source for which it has the (*, G) state.
Step 3	(Optional) **show ipv6 pim group-range** [*ipv6-prefix*	**vrf** *vrf-name*] **Example:** `switch(config)# show ipv6 pim group-range`	Displays PIM6 modes and group ranges.
Step 4	(Optional) **copy running-config startup-config** **Example:** `switch(config-if)# copy running-config startup-config`	Copies the running configuration to the startup configuration.

Configuring SSM

SSM is a multicast distribution mode where the software on the DR connected to a receiver that is requesting data for a multicast source builds a shortest path tree (SPT) to that source.

On an IPv4 network, a host can request multicast data for a specific source only if it is running IGMPv3 and the DR for that host is running IGMPv3. You will usually enable IGMPv3 when you configure an interface for PIM in the SSM mode. For hosts running IGMPv1 or IGMPv2, you can configure group-to-source mapping using SSM translation.

You can only configure the IPv4 group range that is used by SSM.

Note If you want to use the default SSM group range, you do not need to configure the SSM group range.

Before you begin

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

SUMMARY STEPS

1. **configure terminal**
2. [**no**] **ip pim ssm** {**prefix-list** *name* | **range** {*ip-prefix* | none} | route-map *policy-name*}
3. (Optional) **show ip pim group-range** [*ip-prefix* | **vrf** *vrf-name*]
4. (Optional) **copy running-config startup-config**

DETAILED STEPS

Procedure

	Command or Action	Purpose

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

Step 2	[**no**] **ip pim ssm** {**prefix-list** *name* | **range** {*ip-prefix* | none} | route-map *policy-name*}
	Example:
	`switch(config)# ip pim ssm range 239.128.1.0/24`
	Example:
	`switch(config)# no ip pim ssm range none`
		The following options are available:
		• **prefix-list**—Specifies a prefix-list policy name for the SSM range.
		• **range**—Configures a group range for SSM. The default range is 232.0.0.0/8. If the keyword **none** is specified, all group ranges are removed.
		• **route-map**—Specifies a route-map policy name that lists the group prefixes to use with the **match ip multicast** command.
		The **no** option removes the specified prefix from the SSM range or removes the prefix-list or route-map policy. If the keyword **none** is specified, the **no** command resets the SSM range to the default value of 232.0.0.0/8.
		Note
		You can configure a maximum of four ranges for SSM multicast, using the **prefix-list**, **range**, or **route-map** commands.

Step 3	(Optional) **show ip pim group-range** [*ip-prefix* | **vrf** *vrf-name*]	Displays PIM modes and group ranges.

	Command or Action	Purpose
	Example: `switch(config)# show ip pim group-range`	
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 PIM SSM Over a vPC

Configuring PIM SSM over a vPC enables support for IGMPv3 joins and PIM S,G joins over vPC peers in the SSM range. This configuration is supported for orphan sources or receivers in the Layer 2 or Layer 3 domain. When you configure PIM SSM over a vPC, no rendezvous point (RP) configuration is required.

(S,G) entries will have the RPF as the interface toward the source, and no *,G states will be maintained in the MRIB.

Before you begin

Ensure that you have the PIM and vPC features enabled.

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

SUMMARY STEPS

1. **configure terminal**
2. **vrf context** *name*
3. (Optional) [**no**] **ip pim ssm** {**prefix-list** *name* | **range** {*ip-prefix* | **none**} | **route-map** *policy-name*}
4. (Optional) **show ip pim group-range** [*ip-prefix*] [**vrf** *vrf-name* | **all**]
5. (Optional) **copy running-config startup-config**

DETAILED STEPS

Procedure

	Command or Action	Purpose

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

Step 2	**vrf context** *name*
	Example:
	`switch(config)# vrf context Enterprise`
	`switch(config-vrf)#`
		Creates a new VRF and enters VRF configuration mode. The *name* can be any case-sensitive, alphanumeric string up to 32 characters.

	Command or Action	Purpose
Step 3	(Optional) [**no**] **ip pim ssm** {**prefix-list** *name* | **range** {*ip-prefix* | **none**} | **route-map** *policy-name*} **Example:** `switch(config-vrf)# ip pim ssm range 234.0.0.0/24`	The following options are available: • **prefix-list**—Specifies a prefix-list policy name for the SSM range. • **range**—Configures a group range for SSM. The default range is 232.0.0.0/8. If the keyword **none** is specified, all group ranges are removed. • **route-map**—Specifies a route-map policy name that lists the group prefixes to use with the **match ip multicast** command. By default, the SSM range is 232.0.0.0/8. PIM SSM over vPC works as long as S,G joins are received in this range. If you want to override the default with some other range, you must specify that range using this command. The command in the example overrides the default range to 234.0.0.0/24. The **no** option removes the specified prefix from the SSM range or removes the prefix-list or route-map policy. If the keyword **none** is specified, the **no** command resets the SSM range to the default value of 232.0.0.0/8.
Step 4	(Optional) **show ip pim group-range** [*ip-prefix*] [**vrf** *vrf-name* | **all**] **Example:** `switch(config-vrf)# show ip pim group-range`	Displays PIM modes and group ranges.
Step 5	(Optional) **copy running-config startup-config** **Example:** `switch(config-vrf)# copy running-config startup-config`	Copies the running configuration to the startup configuration.

Configuring RPF Routes for Multicast

You can define reverse path forwarding (RPF) routes for multicast when you want multicast data to diverge from the unicast traffic path. You can define RPF routes for multicast on border routers to enable RPF to an external network.

Multicast routes are used not to directly forward traffic but to make RPF checks. RPF routes for multicast cannot be redistributed.

Note IPv6 static multicast routes are not supported.

Note If the **ip multicast multipath s-g-hash** CLI is not configured, the multicast traffic may fail the RFP check.

Before you begin

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

SUMMARY STEPS

1. **configure terminal**
2. **ip mroute** {*ip-addr mask* | *ip-prefix*} {*next-hop* | *nh-prefix* | *interface*} [*route-preference*] [**vrf** *vrf-name*]
3. (Optional) **show ip static-route** [**multicast**] [**vrf** *vrf-name*]
4. (Optional) **copy running-config startup-config**

DETAILED STEPS

Procedure

	Command or Action	Purpose

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

Step 2	**ip mroute** {*ip-addr mask* | *ip-prefix*} {*next-hop* | *nh-prefix* | *interface*} [*route-preference*] [**vrf** *vrf-name*] **Example:** `switch(config)# ip mroute 192.0.2.33/1 224.0.0.0/1`	Configures an RPF route for multicast for use in RPF calculations. Route preference values range from 1 to 255. The default preference is 1.
Step 3	(Optional) **show ip static-route** [**multicast**] [**vrf** *vrf-name*] **Example:** `switch(config)# show ip static-route multicast`	Displays configured static routes.
Step 4	(Optional) **copy running-config startup-config**	Copies the running configuration to the startup configuration.

Configuring Multicast Multipath

By default, the RPF interface for multicast is chosen automatically when multiple ECMP paths are available.

SUMMARY STEPS

1. **configure terminal**
2. **ip multicast multipath** {**none** | **resilient** | **s-g-hash**}
3. **clear ip mroute** *

DETAILED STEPS

Procedure

	Command or Action	Purpose

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

Step 2	**ip multicast multipath** {**none** | **resilient** | **s-g-hash**} **Example:** `switch(config)# ip multicast multipath none`	Configure multicast multipath using the following options: • **none**—Disables multicast multipath by suppressing hashing across multiple ECMPs in the URIB RPF lookup. With this option, the highest RPF neighbor (next-hop) address is used for the RPF interface. **Note** Use the **ip multicast multipath none** command to completely disable hashing. • **s-g-hash**—Initiates S, G, nexthop hashing (rather than the default of S/RP, G-based hashing) to select the RPF interface. This option configures the hash based on source and group address. This is the default setting. • **resilient**—If the ECMP path list changes and the old RPF information is still part of the ECMP, this option uses the old RPF information instead of performing a rehash and potentially changing the RPF information. The **ip multicast multipath resilient** command is for maintaining resiliency (Stickiness) to the current RPF if there is a path in the route reachability notification from URIB. **Note** The **no ip multicast multipath resilient** command disables the stickiness algorithm. This command is independent of the hashing algorithm.
Step 3	**clear ip mroute *** **Example:** `switch(config)# clear ip mroute *`	Clears multipath routes and activates multicast multipath suppression.

Configuring Multicast VRF-Lite Route Leaking

Beginning with Cisco NX-OS Release 7.0(3)I7(1), you can configure multicast VRF-lite route leaking, which allows IPv4 multicast traffic across VRFs.

Before you begin

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

SUMMARY STEPS

1. **configure terminal**
2. **ip multicast rpf select vrf** *src-vrf-name* **group-list** *group-list*
3. (Optional) **copy running-config startup-config**

DETAILED STEPS

Procedure

	Command or Action	Purpose

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

Step 2	**ip multicast rpf select vrf** *src-vrf-name* **group-list** *group-list* **Example:** `switch(config)# ip multicast rpf select vrf blue group-list 236.1.0.0/16`	Specifies which VRF to use for RPF lookup for a particular multicast group. *src-vrf-name* is the name of the source VRF. It can be a maximum of 32 alphanumeric characters and is case sensitive. *group-list* is the group range for the RPF. The format is A.B.C.D/LEN with a maximum length of 32.
Step 3	(Optional) **copy running-config startup-config** **Example:** `switch(config)# copy running-config startup-config`	Copies the running configuration to the startup configuration.

Configuring Route Maps to Control RP Information Distribution

You can configure route maps to help protect against some RP configuration errors and malicious attacks.

By configuring route maps, you can control distribution of RP information that is distributed throughout the network. You specify the BSRs or mapping agents to be listened to on each client router and the list of candidate RPs to be advertised (listened to) on each BSR and mapping agent to ensure that what is advertised is what you expect.

Note Only the **match ipv6 multicast** command has an effect in the route map.

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

Configuring Route Maps to Control RP Information Distribution (PIM)

SUMMARY STEPS

1. **configure terminal**
2. **route-map** *map-name* [**permit** | **deny**] [*sequence-number*]
3. **match ip multicast** {**rp** *ip-address* [**rp-type** *rp-type*]} {**group** *ip-prefix*} {**source** *source-ip-address*}
4. (Optional) **show route-map**
5. (Optional) **copy running-config startup-config**

DETAILED STEPS

Procedure

	Command or Action	Purpose

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

Step 2	**route-map** *map-name* [**permit**	**deny**] [*sequence-number*]
		Example:
		`switch(config)# route-map ASM_only permit 10`
		`switch(config-route-map)#`
			Enters route-map configuration mode.

Step 3	**match ip multicast** {**rp** *ip-address* [**rp-type** *rp-type*]} {**group** *ip-prefix*} {**source** *source-ip-address*} **Example:** `switch(config-route-map)# match ip multicast group 224.0.0.0/4 rp 0.0.0.0/0 rp-type ASM`	Matches the group, RP, and RP type specified. You can specify the RP type (ASM). This configuration method requires the group and RP specified as shown in the example. **Note** In the **match ip multicast group-range** <> CLI, the **group-range** command under **route-map** is not supported.
Step 4	(Optional) **show route-map** **Example:** `switch(config-route-map)# show route-map`	Displays configured route maps.
Step 5	(Optional) **copy running-config startup-config** **Example:** `switch(config-route-map)# copy running-config startup-config`	Copies the running configuration to the startup configuration.

Configuring Route Maps to Control RP Information Distribution (PIM6)

SUMMARY STEPS

1. **configure terminal**
2. **route-map** *map-name* [**permit** | **deny**] [*sequence-number*]
3. **match ipv6 multicast** {**rp** *ip-address* [**rp-type** *rp-type*]} {**group** *ipv6-prefix*} {**source** *source-ip-address*}
4. (Optional) **show route-map**
5. (Optional) **copy running-config startup-config**

DETAILED STEPS

Procedure

	Command or Action	Purpose

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

Step 2	**route-map** *map-name* [**permit**	**deny**] [*sequence-number*]
		Example:
		`switch(config)# route-map ASM_only permit 10`
		`switch(config-route-map)#`
			Enters route-map configuration mode.

Step 3	**match ipv6 multicast** {**rp** *ip-address* [**rp-type** *rp-type*]} {**group** *ipv6-prefix*} {**source** *source-ip-address*} **Example:** `switch(config-route-map)# match ipv6 multicast group ff1e:abcd:def1::0/24 rp 2001:0db8:0:abcd::1 rp-type ASM`	Matches the group, RP, and RP type specified. You can specify the RP type (ASM). This configuration method requires the group and RP specified as shown in the example.
Step 4	(Optional) **show route-map** **Example:** `switch(config-route-map)# show route-map`	Displays configured route maps.
Step 5	(Optional) **copy running-config startup-config** **Example:** `switch(config-route-map)# copy running-config startup-config`	Copies the running configuration to the startup configuration.

Configuring Message Filtering

Note Prefix matches in the rp-candidate-policy must be exact relative to what the c-rp is advertising. Subset matches are not possible.

You can configure filtering of the PIM messages described in the table below.

Table 17: PIM Message Filtering

Message Type	Description
Global to the Device	
Log Neighbor changes	Enables syslog messages that list the neighbor state changes to be generated. The default is disabled.
PIM register policy	Enables PIM register messages to be filtered based on a route-map policy[4] where you can specify group or group and source addresses with the **match ip multicast** command. This policy applies to routers that act as an RP. The default is disabled, which means that the software does not filter PIM register messages.
BSR candidate RP policy	Enables BSR candidate RP messages to be filtered by the router based on a route-map policy where you can specify the RP and group addresses with the **match ip multicast** command. This command can be used on routers that are eligible for BSR election. The default is no filtering of BSR messages.
BSR policy	Enables BSR messages to be filtered by the BSR client routers based on a route-map policy where you can specify BSR source addresses with the **match ip multicast** command. This command can be used on client routers that listen to BSR messages. The default is no filtering of BSR messages.
Auto-RP candidate RP policy	Enables Auto-RP announce messages to be filtered by the Auto-RP mapping agents based on a route-map policy where you can specify the RP and group addresses with the **match ip multicast** command. This command can be used on a mapping agent. The default is no filtering of Auto-RP messages.
Auto-RP mapping agent policy	Enables Auto-RP discover messages to be filtered by client routers based on a route-map policy where you can specify mapping agent source addresses with the **match ip multicast** command. This command can be used on client routers that listen to discover messages. The default is no filtering of Auto-RP messages. **Note** PIM6 does not support the Auto-RP method.
Per Device Interface	

Message Type	Description
Join-prune policy	Enables join-prune messages to be filtered based on a route-map policy where you can specify group, group and source, or group and RP addresses with the **match ip multicast** command. The default is no filtering of join-prune messages.

[4] For information about configuring route-map policies, see the *Cisco Nexus 9000 Series NX-OS Unicast Routing Configuration Guide*.

Route maps as a filtering policy can be used (either **permit** or **deny** for each statement) for the following commands:

- The **jp-policy** command can use (S,G), (*,G), or (RP,G).
- The **register-policy** command can use (S,G) or (*,G).
- The **igmp report-policy** command can use (*,G) or (S,G).
- The **state-limit reserver-policy** command can use (*,G) or (S,G).
- The **auto-rp rp-candidate-policy** command can use (RP,G).
- The **bsr rp-candidate-policy** command can use (RP,G).
- The **autorp mapping-agent policy** command can use (S).
- The **bsr bsr-policy** command can use (S).

Route maps as containers can be used for the following commands, where the route-map action (**permit** or **deny**) is ignored:

- The **ip pim rp-address route map** command can use only G.
- The **ip igmp static-oif route map** command can use (S,G), (*,G), (S,G-range), (*,G-range).
- The **ip igmp join-group route map** command can use (S,G), (*,G), (S,G-range, (*, G-range).

Configuring Message Filtering

Before you begin

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

SUMMARY STEPS

1. **configure terminal**
2. (Optional) **ip pim log-neighbor-changes**
3. (Optional) **ip pim register-policy** *policy-name*
4. (Optional) **ip pim bsr rp-candidate-policy** *policy-name*
5. (Optional) **ip pim bsr bsr-policy** *policy-name*
6. (Optional) **ip pim auto-rp rp-candidate-policy** *policy-name*
7. (Optional) **ip pim auto-rp mapping-agent-policy** *policy-name*
8. **interface** *interface*

9. (Optional) **ip pim jp-policy** *policy-name* [**in** | **out**]
10. (Optional) **show run pim**
11. (Optional) **copy running-config startup-config**

DETAILED STEPS

Procedure

	Command or Action	Purpose

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

Step 2	(Optional) **ip pim log-neighbor-changes** **Example:** `switch(config)# ip pim log-neighbor-changes`	Enables syslog messages that list the neighbor state changes to be generated. The default is disabled.
Step 3	(Optional) **ip pim register-policy** *policy-name* **Example:** `switch(config)# ip pim register-policy my_register_policy`	Enables PIM register messages to be filtered based on a route-map policy. You can specify group or group and source addresses with the **match ip multicast** command.
Step 4	(Optional) **ip pim bsr rp-candidate-policy** *policy-name* **Example:** `switch(config)# ip pim bsr rp-candidate-policy my_bsr_rp_candidate_policy`	Enables BSR candidate RP messages to be filtered by the router based on a route-map policy where you can specify the RP and group addresses with the **match ip multicast** command. This command can be used on routers that are eligible for BSR election. The default is no filtering of BSR messages.
Step 5	(Optional) **ip pim bsr bsr-policy** *policy-name* **Example:** `switch(config)# ip pim bsr bsr-policy my_bsr_policy`	Enables BSR messages to be filtered by the BSR client routers based on a route-map policy where you can specify BSR source addresses with the **match ip multicast** command. This command can be used on client routers that listen to BSR messages. The default is no filtering of BSR messages.
Step 6	(Optional) **ip pim auto-rp rp-candidate-policy** *policy-name* **Example:** `switch(config)# ip pim auto-rp rp-candidate-policy my_auto_rp_candidate_policy`	Enables Auto-RP announce messages to be filtered by the Auto-RP mapping agents based on a route-map policy where you can specify the RP and group addresses with the **match ip multicast** command. This command can be used on a mapping agent. The default is no filtering of Auto-RP messages.
Step 7	(Optional) **ip pim auto-rp mapping-agent-policy** *policy-name* **Example:** `switch(config)# ip pim auto-rp mapping-agent-policy my_auto_rp_mapping_policy`	Enables Auto-RP discover messages to be filtered by client routers based on a route-map policy where you can specify mapping agent source addresses with the **match ip multicast** command. This command can be used on client

	Command or Action	Purpose
		routers that listen to discover messages. The default is no filtering of Auto-RP messages.

Step 8	**interface** *interface*
	Example:
	`switch(config)# interface ethernet 2/1`
	`switch(config-if)#`
		Enters interface mode on the specified interface.

Step 9	(Optional) **ip pim jp-policy** *policy-name* [**in** | **out**] **Example:** `switch(config-if)# ip pim jp-policy my_jp_policy`	Enables join-prune messages to be filtered based on a route-map policy where you can specify group, group and source, or group and RP addresses with the **match ip multicast** command. The default is no filtering of join-prune messages.
Step 10	(Optional) **show run pim** **Example:** `switch(config-if)# show run pim`	Displays PIM configuration commands.
Step 11	(Optional) **copy running-config startup-config** **Example:** `switch(config-if)# copy running-config startup-config`	Copies the running configuration to the startup configuration.

Configuring Message Filtering (PIM6)

Before you begin

Ensure that you have installed the Enterprise Services license and enabled PIM6.

SUMMARY STEPS

1. **configure terminal**
2. (Optional) **ipv6 pim log-neighbor-changes**
3. (Optional) **ipv6 pim register-policy** *policy-name*
4. (Optional) **ipv6 pim jp-policy** *policy-name* [**in** | **out**]
5. (Optional) **show run pim6**
6. (Optional) **copy running-config startup-config**

DETAILED STEPS

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	(Optional) **ipv6 pim log-neighbor-changes** **Example:** `switch(config)# ipv6 pim log-neighbor-changes`	Enables syslog messages that list the neighbor state changes to be generated. The default is disabled.

Step 3	(Optional) **ipv6 pim register-policy** *policy-name*
	Example:
	`switch(config)# ipv6 pim register-policy my_register_policyinterface interfaceEnters interface mode on the specified interface.`
	`switch(config)# interface ethernet 2/1`
	`switch(config-if)#`
		Enables PIM register messages to be filtered based on a route-map policy. You can specify group or group and source addresses with the **match ipv6 multicast** command. The default is disabled.

Step 4	(Optional) **ipv6 pim jp-policy** *policy-name* [**in** | **out**] **Example:** `switch(config-if)# ipv6 pim jp-policy my_jp_policy`	Enables join-prune messages to be filtered based on a route-map policy where you can specify group, group and source, or group and RP addresses with the **match ipv6 multicast** command. The default is no filtering of join-prune messages. This command filters messages in both incoming and outgoing directions.
Step 5	(Optional) **show run pim6** **Example:** `switch(config-if)# show run pim6`	Displays PIM6 configuration commands.
Step 6	(Optional) **copy running-config startup-config** **Example:** `switch(config-if)# copy running-config startup-config`	Copies the running configuration to the startup configuration.

Restarting the PIM Processes

When routes are flushed, they are removed from the Multicast Routing Information Base (MRIB) and the Multicast Forwarding Information Base (MFIB).

When you restart PIM, the following tasks are performed:

- The PIM database is deleted.
- The MRIB and MFIB are unaffected and forwarding of traffic continues.
- The multicast route ownership is verified through the MRIB.
- Periodic PIM join and prune messages from neighbors are used to repopulate the database.

Restarting the PIM Process

Before you begin

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

SUMMARY STEPS

1. **restart pim**
2. **configure terminal**
3. **ip pim flush-routes**
4. (Optional) **show running-configuration pim**
5. (Optional) **copy running-config startup-config**

DETAILED STEPS

Procedure

	Command or Action	Purpose
Step 1	**restart pim** **Example:** `switch# restart pim`	Restarts the PIM process. **Note** Traffic loss might occur during the restart process.

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

Step 3	**ip pim flush-routes** **Example:** `switch(config)# ip pim flush-routes`	Removes routes when the PIM process is restarted. By default, routes are not flushed.
Step 4	(Optional) **show running-configuration pim** **Example:** `switch(config)# show running-configuration pim`	Displays the PIM running-configuration information, including the **flush-routes** command.
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 PIM6 Process

Before you begin

Ensure that you have installed the Enterprise Services license and enabled PIM6.

SUMMARY STEPS

1. **restart pim6**
2. **configure terminal**
3. **ipv6 pim flush-routes**
4. (Optional) **show running-configuration pim6**
5. (Optional) **copy running-config startup-config**

DETAILED STEPS

Procedure

	Command or Action	Purpose
Step 1	**restart pim6** **Example:** `switch# restart pim6`	Restarts the PIM6 process.

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

Step 3	**ipv6 pim flush-routes** **Example:** `switch(config)# ipv6 pim flush-routes`	Removes routes when the PIM6 process is restarted. By default, routes are not flushed.
Step 4	(Optional) **show running-configuration pim6** **Example:** `switch(config)# show running-configuration pim6`	Displays the PIM6 running-configuration information, including the **flush-routes** command.
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 BFD for PIM in VRF Mode

Note You can configure Bidirectional Forwarding Detection (BFD) for PIM by either VRF or interface.

Before you begin

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

SUMMARY STEPS

1. **configure terminal**
2. **vrf context** *vrf-name*
3. **ip pim bfd**

DETAILED STEPS

Procedure

	Command or Action	Purpose

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

Step 2	**vrf context** *vrf-name*
	Example:
	`switch# vrf context test`
	`switch(config-vrf)#`
		Enters VRF configuration mode.

Step 3	**ip pim bfd** **Example:** `switch(config-vrf)# ip pim bfd`	Enables BFD on the specified VRF. **Note** You can also enter the **ip pim bfd** command in global configuration mode, which enables BFD on the VRF instance.

Configuring BFD for PIM in Interface Mode

Before you begin

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

SUMMARY STEPS

1. **configure terminal**
2. **interface** *interface-type*
3. **ip pim bfd instance**
4. (Optional) **show running-configuration pim**
5. (Optional) **copy running-config startup-config**

DETAILED STEPS

Procedure

	Command or Action	Purpose

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

Step 2	**interface** *interface-type*
	Example:
	`switch(config)# interface ethernet 7/40`
	`switch(config-if)#`
		Enters interface configuration mode.

Step 3	**ip pim bfd instance** **Example:** `switch(config-if)# ip pim bfd instance`	Enables BFD on the specified interfaces. You can enable or disable BFD on PIM interfaces irrespective of whether BFD is enabled on the VRF.
Step 4	(Optional) **show running-configuration pim** **Example:** `switch(config-if)# show running-configuration pim`	Displays the PIM running-configuration information.
Step 5	(Optional) **copy running-config startup-config** **Example:** `switch(config-if)# copy running-config startup-config`	Copies the running configuration to the startup configuration.

Enabling the Multicast Heavy and Extended Heavy Template

You can enable the multicast heavy template in order to support up to 32K IPv4 mroutes.

You must enable the multicast extended heavy template and configure the multicast route memory in order to support the 128K IPv4 route.

With the heavy template, the **show ip mroute** command displays the multicast traffic counters.

Before you begin

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

SUMMARY STEPS

1. **configure terminal**
2. **system routing** *template-name*
3. **vdc** *vdc-name*
4. **limit-resource m4route-mem** [**minimum** *min-value*]**maximum** *max-value*
5. **exit**

6. **ip routing multicast mfdm-buffer-route-count** *size*
7. **ip pim mtu** *size*
8. **exit**
9. **show system routing mode**
10. (Optional) **copy running-config startup-config**

DETAILED STEPS

Procedure

	Command or Action	Purpose

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

Step 2	**system routing** *template-name*
	Example:
	`switch(config)# system routing template-multicast-heavy`
	`switch(config)# system routing template-multicast-ext-heavy`
	`switch(config)# system routing template-dual-stack-mcast`
		Enables the multicast template. The template can be template-multicast-heavy or template-multicast-ext-heavy or template-dual-stack-mcast. You need to reload the system after enabling the command when you use template-multicast-heavy or template-multicast-ext-heavy templates.

Step 3	**vdc** *vdc-name* **Example:** `switch(config)# vdc vdc1`	Specifies a VDC and enters VDC configuration mode.
Step 4	**limit-resource m4route-mem** [**minimum** *min-value*]**maximum** *max-value* **Example:** `switch(config-vdc)# limit-resource m4route-mem minimum 150 maximum 150`	Configures IPv4 multicast route map memory resource limits for a VDC. After configuring this command, save it to the startup configuration and reload the device.
Step 5	**exit** **Example:** `switch(config-vdc)# exit`	Exits VDC configuration mode.
Step 6	**ip routing multicast mfdm-buffer-route-count** *size* **Example:** `switch(config)# ip routing multicast mfdm-buffer-route-count 400`	Configures the multicast mfdm buffer route size.
Step 7	**ip pim mtu** *size* **Example:** `switch(config)# ip pim mtu 1500`	Enables bigger frame sizes for the PIM control plane traffic and improves the convergence.

	Command or Action	Purpose
Step 8	**exit** **Example:** `switch(config)# exit`	Exits the global configuration mode.

Step 9	**show system routing mode**
	Example:
	`switch# show system routing mode`
	`Configured System Routing Mode: Multicast Extended Heavy Scale`
	`Applied System Routing Mode: Multicast Extended Heavy Scale`
	`Switch#`
		Displays the configured routing mode - multicast heavy or multicast extended heavy or dual stack.

Step 10	(Optional) **copy running-config startup-config** **Example:** `switch(config)# copy running-config startup-config`	Copies the running configuration to the startup configuration.

Verifying the PIM Configuration

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

Command	Description
show ip mroute [*ip-address*] [**detail** | **summary**]	Displays the IP multicast routing table. The **detail** option displays detailed route attributes. The **summary** option displays route counts and packet rates.
show ip pim group-range [*ip-prefix*] [**vrf** *vrf-name* | **all**]	Displays the learned or configured group ranges and modes. For similar information, see the **show ip pim rp** command.
show ip pim interface [*interface* | **brief**] [**vrf** *vrf-name* | **all**]	Displays information by the interface.
show ip pim neighbor [**interface** *interface* | *ip-prefix*] [**vrf** *vrf-name* | **all**]	Displays neighbors by the interface.
show ip pim oif-list *group* [*source*] [**vrf** *vrf-name* | **all**]	Displays all the interfaces in the outgoing interface (OIF) list.

Command	Description
show ip pim route [*source* | *group* [*source*]] [**vrf** *vrf-name* | **all**]	Displays information for each multicast route, including interfaces on which a PIM join for that (S, G) has been received.
show ip pim rp [*ip-prefix*] [**vrf** *vrf-name* | **all**]	Displays rendezvous points (RPs) known to the software, how they were learned, and their group ranges. For similar information, see the **show ip pim group-range** command.
show ip pim rp-hash *group* [**vrf** *vrf-name* | **all**]	Displays the bootstrap router (BSR) RP hash information.
show running-config pim	Displays the running-configuration information.
show startup-config pim	Displays the startup-configuration information.
show ip pim vrf [*vrf-name* | **all**] [**detail**]	Displays per-VRF information.

Displaying Statistics

You can display and clear PIM statistics by using the commands in this section.

Displaying PIM Statistics

You can display the PIM statistics and memory usage using these commands.

Command	Description
show ip pim policy statistics	Displays policy statistics for register, RP, and join-prune message policies.
show ip pim statistics [**vrf** *vrf-name*]	Displays global statistics.

Clearing PIM Statistics

You can clear the PIM statistics using these commands.

Command	Description
clear ippim interface statistics *interface*	Clears counters for the specified interface.
clear ip pim policy statistics	Clears policy counters for register, RP, and join-prune message policies.
clear ip pim statistics [**vrf** *vrf-name*]	Clears global counters handled by the PIM process.

Configuring Multicast Service Reflection

The Multicast Service Reflection feature enables you to translate externally received multicast destination addresses to addresses that conform to your organization's internal addressing policy. It is the multicast Network Address Translation (NAT) of an externally received multicast stream (S1,G1) to (S2,G2) into the internal domain. Unlike IP NAT, which only translates the source IP address, the Multicast Service Reflection translates both the source and destination addresses.

The Ingress NAT allows translation of incoming (S,G) into a different source, group or both. All receivers inside the domain then can join the post translated flow. This feature is useful when multicast traffic:

- enters a network from a different domain with potentially overlapping address
- comes with an address that is not understood by applications in the network

The Egress NAT allows translating existing flow (S,G) to different source or group address on a per outgoing interface basis. This feature is useful for multicast distribution to external entities which may only accept a certain source, group address. It can also serve as a way to hide internal address space when flows are exposed to external entities.

The Multicast Service Reflection feature is configured on a loopback interface in the VRF configuration mode. The flow incoming as S1, G1 is translated to S2, G2 and the destination MAC address is re-written to the multicast MAC address of translated address which is G2.

Guidelines and Limitations for Multicast Service Reflection

The Multicast Service Reflection feature has the following guidelines and limitations:

- The Multicast Service Reflection feature is introduced in Cisco NX-OS Release 9.3(5) and it is is supported on the Cisco Nexus 9300-FX, FX2, FXP, EX Series switches.
- Beginning with Cisco NX-OS Release 9.3(5)F, the range for maximum replications for the map interface is 1-40.
- Beginning with Cisco NX-OS Release 9.3(3)F, multicast NAT is supported on Cisco Nexus C9300-GX.
- Beginning with Cisco NX-OS Release 10.1(1)F, Multicast Service Reflection with NBM is supported on Cisco Nexus 9300-FX3, Cisco Nexus C9316D-GX, Cisco Nexus C93600CD-GX, and Cisco Nexus C9364C-GX platform switches.
- The Multicast Service Reflection feature is not supported on the following platforms:
  - Cisco Nexus 9500 series switches with cloud scale line cards
  - Cisco Nexus 9500 series switches with R-series line cards

  - Cisco Nexus 3600-R series switches
  - Cisco Nexus 9200 series switches
  - Cisco Nexus 9364C switches

- The Multicast Service Reflection feature is supported in Protocol Independent Multicast (PIM) sparse mode only (ASM or SSM).
- The Multicast Service Reflection feature does not work in a vPC environment.
- Multicast-to-Unicast translation is not supported in Cisco NX-OS Release 10.1(x).
- Multicast-to-Multicast and Unicast-to-Unicast NAT configuration cannot be done together and at the same time.
- Unicast NAT, Multicast NAT, and PBR features are not supported at the same time on the same device.
- Egress NAT functionality is supported only under default VRF and not under other VRFs.
- FEX is not supported.
- Multicast Service Reflection feature does not support non-NATed receivers for pre-translated (S1,G1) if a NAT rule is configured for this pair (i.e., ingress NAT does not support the pre-translated (S1,G1) receivers while the egress NAT supports them). The untranslated receiver OIFs are supported for egress NAT.
- SVI is not supported for RPF and OIFs.
- Subinterface receiver for post-translated Egress NAT groups is not supported.
- The selected hardware loopback port for a Multicast Service Reflection configuration should be a physical port with a 'Link Down' state and with no SFP connected.
- The multicast NAT translation does not happen with the mask length 0 to 4. This mask length limitation is only for the group address and it is not for the source addresses.
- For IGMP static joins on interfaces the group range mask of /24 are used to generate the joins. The source mask length is considered as /32. A variation in source mask length is not considered in generating the joins in the **ip igmp static** join command.

Ingress and egress interface ACLs on a device configured for the Multicast Service Reflection feature have the following limitations:

- When an ingress ACL is applied to block the untranslated multicast traffic that is already flowing, the (S,G) entries are not removed. The reason is that the multicast route entries continue to be hit by the traffic, even though the ACL drops the packets.
- When an egress ACL is applied to block translated source traffic (S2,G2) on an egress interface, the egress ACL does not work because an egress ACL is not supported for the translated traffic.

Prerequisites

Multicast Service Reflection feature has the following prerequisite:

For platforms that support the Multicast Service Reflection feature, TCAM needs to be carved before configuring Multicast NAT. Use the following command:

```
hardware access-list tcam region mcast-nat region tcam-size
```

Configuring Multicast Service Reflection

Before you begin

- Make sure your multicast-enabled network runs either Protocol Independent Multicast Sparse Mode (PIM-SM) or PIM Source Specific Multicast (PIM-SSM).
- Make sure that the virtual interface for Multicast Service Reflection is configured in your NAT router and the Multicast Service Reflection rules are configured and operational.

SUMMARY STEPS

1. **configure terminal**
2. **vrf context** *name*
3. **[no] ip service-reflect source-interface** *interface-name interface-number*
4. **[no] ip service-reflect destination** *in-grp* **to** *out-grp* **mask-len** *g-mlen* **source** *in-src* **to** *out-src* **mask-len** *s-mlen*[**to-udp** *udp-to-src-port udp-to-dest-port*] [**to-udp-src-port** *udp-to-src-port*] [**to-udp-dest-port** *udp-to-dest-port*]
5. **[no] ip service-reflect mode egress** *prefix*
6. **[no] ip service-reflect destination** *in-grp* **to** *out-grp* **mask-len** *g-mlen* **source** *in-src* **to** *out-src* **mask-len** *s-mlen*[**to-udp** *udp-to-src-port udp-to-dest-port*] [**to-udp-src-port** *udp-to-src-port*] [**to-udp-dest-port** *udp-to-dest-port*] [**static-oif** *out-if*]
7. **[no] multicast service-reflect interface all map interface** *interface-name* **max-replication** *replication*
8. **exit**
9. **interface** *interface-name interface-number*
10. **ip address** *prefix*
11. **ip pim sparse-mode**
12. **ip igmp static-oif** {*group* [**source** *source*] |**route-map** *policy-name*}
13. **no system multicast dcs-check**
14. **ip pim border-router**
15. **nbm external-link**
16. **exit**
17. **[no] multicast service-reflect interface all map interface** *interface-name* **vrf** *vrf-name*
18. **[no] multicast service-reflect interface** *interface-name* **map interface** *interface-name***vrf** *vrf-name*
19. **[no] multicast service-reflect interface** *interface-1, interface-2, interface-3***map interface** *interface-name***vrf** *vrf-name*
20. **exit**
21. **show ip mroute sr**
22. **show forwarding distribution multicast route**
23. **show forwarding distribution multicast route group**

DETAILED STEPS

Procedure

	Command or Action	Purpose

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

Step 2	**vrf context** *name*
	Example:
	`switch(config)# vrf context test`
	`switch(config-vrf)#`
		Creates a new VRF and enters VRF configuration mode. The *name* can be any case-sensitive, alphanumeric string up to 32 characters. The NAT rules are configured the vrf context.
		Note
		Non-default VRF is not supported for egress NAT.

Step 3	**[no] ip service-reflect source-interface** *interface-name interface-number* **Example:** `switch(config-vrf)# ip service-reflect source-interface loopback10`	Configures a loopback as the NAT source. This interface pulls traffic to the NAT router. The interface will be RPF for the post translated routes. This command is configured per VRF.
Step 4	**[no] ip service-reflect destination** *in-grp* **to** *out-grp* **mask-len** *g-mlen* **source** *in-src* **to** *out-src* **mask-len** *s-mlen*[**to-udp** *udp-to-src-port udp-to-dest-port*] [**to-udp-src-port** *udp-to-src-port*] [**to-udp-dest-port** *udp-to-dest-port*] **Example:** `switch(config-vrf)# ip service-reflect destination 228.1.1.1 to 238.1.1.1 mask-len 32 source 80.80.80.80 to 90.90.90.90 mask-len 32 to-udp-src-port 500 to-udp-dest-port 600`	Configures the NAT rule for the ingress NAT.
Step 5	**[no] ip service-reflect mode egress** *prefix* **Example:** `switch(config-vrf)# ip service-reflect mode egress 225.1.1.0/24`	Configures the egress NAT mode. Matches and rewrites multicast packets routed on to the interface. **Note** Egress NAT is supported only on the default VRF.
Step 6	**[no] ip service-reflect destination** *in-grp* **to** *out-grp* **mask-len** *g-mlen* **source** *in-src* **to** *out-src* **mask-len** *s-mlen*[**to-udp** *udp-to-src-port udp-to-dest-port*] [**to-udp-src-port** *udp-to-src-port*] [**to-udp-dest-port** *udp-to-dest-port*] [**static-oif** *out-if*] **Example:** `switch(config-vrf)# ip service-reflect destination 225.1.1.1 to 227.1.1.1 mask-len 32 source 10.10.10.100 to 20.10.10.101 mask-len 32`	Configures the NAT rule for the egress NAT.

	Command or Action	Purpose
	`to-udp-src-port 33 to-udp-dest-port 66 static-oif Ethernet1/8`	
Step 7	**[no] multicast service-reflect interface all map interface** *interface-name* **max-replication** *replication* **Example:** `switch(config-vrf)# multicast service-reflect interface all map interface Ethernet1/54 max-replication 3`	Specifies the maximum replications for the map interface. The range is 1–40. Default value is 40. The **no** command deletes the configuration.

Step 8	**exit**
	Example:
	`switch(config-vrf)# exit`
	`switch(config)#`
		Exits the VRF configuration mode and enters the global configuration mode.

Step 9	**interface** *interface-name interface-number*
	Example:
	`switch(config)# interface loopback10`
	`switch(config-if)#`
		Enters interface configuration mode.

Step 10	**ip address** *prefix* **Example:** `switch(config-if)# ip address 1.1.1.1/24`	Configures an IP address for the loopback interface. It should be a unique IP address that helps to identify this router.
Step 11	**ip pim sparse-mode** **Example:** `switch(config-if)# ip pim sparse-mode`	Enables PIM sparse mode on the interface. The default is disabled.
Step 12	**ip igmp static-oif** {*group* [**source** *source*] |**route-map** *policy-name*} **Example:** `switch(config-if)# ip igmp static-oif 230.1.1.1`	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. Enables the configured loopback interface to join the multicast stream that is to be NATed.
Step 13	**no system multicast dcs-check** **Example:** `switch(config-if)# no system multicast dcs-check`	Enables multicast packets punt to CPU on non-FHR devices for route learning. This is generally used when **ip pim border-router** or **ip igmp host-proxy** features are enabled. This command is not supported on the Cisco Nexus 9300 series and Cisco Nexus 9200 series EOR switches, Cisco Nexus 9504 and Cisco Nexus 9508 EOR and TOR switches, and N3K-C3636C-R, N3K-C36180YC-R TOR switches.

	Command or Action	Purpose
Step 14	**ip pim border-router** **Example:** `switch(config-if)# ip pim border-router`	Ensures that the traffic from sources outside the PIM-SM domain reaches the receivers inside the domain and allows the remotely sourced traffic to reach local receivers in this domain. A PIM Border Router is required when no PIM messages can traverse the PIM domain border.
Step 15	**nbm external-link** **Example:** `switch(config-if)# nbm external-link`	Configures the NBM interface as an external link in order to connect multiple fabrics together in a multisite solution. **Note** This command is needed only if feature NBM is enabled and on the links where the **ip pim border-router** command is enabled.

Step 16	**exit**
	Example:
	`switch(config-if)# exit`
	`switch(config)#`
		Exits the interface configuration mode and enters the global configuration mode.

Step 17	**[no] multicast service-reflect interface all map interface** *interface-name* **vrf** *vrf-name* **Example:** `switch(config)# multicast service-reflect interface all map interface loopback10 vrf test`	Maps all the fan-out interfaces to a service interface. **Note** The **vrf** *vrf-name* option is not supported for egress NAT. **Note** The commands in steps 17, 18, and 19 are needed only in case of Egress NAT. Each OIF used in the Egress NAT rules configuration need to be mapped to one service-interface using one of these mapping configurations.
Step 18	**[no] multicast service-reflect interface** *interface-name* **map interface** *interface-name***vrf** *vrf-name* **Example:** `switch(config)# multicast service-reflect interface ethernet1/18 map interface loopback10 vrf test`	Configures one-to-one mapping of fan-out interface to a service interface.
Step 19	**[no] multicast service-reflect interface** *interface-1*, *interface-2*, *interface-3***map interface** *interface-name***vrf** *vrf-name* **Example:** `switch(config)# multicast service-reflect interface ethernet 1/1-10, ethernet1/12-14, ethernet1/16 map interface loopback10 vrf test`	Configures multi-to-one mapping of fan-out interfaces to a service interface.
Step 20	**exit** **Example:**	Exits the global configuration mode and enters the privileged EXEC mode.

	Command or Action	Purpose
	`switch(config)# exit`	
Step 21	**show ip mroute sr** **Example:** `switch# show ip mroute sr`	Displays the service reflection mroute entries.
Step 22	**show forwarding distribution multicast route** **Example:** `switch# show forwarding distribution multicast route`	Displays information about the pre-translated and the post-translated route information for the egress NAT and pre-translated route information for the ingress NAT.
Step 23	**show forwarding distribution multicast route group** **Example:** `switch# show forwarding distribution multicast route group`	Displays information about the multicast FIB distribution IPv4 multicast routes.

Configuration Examples for Multicast Service Reflection

The following example shows the Multicast NAT - ingress and egress configuration:

```
interface loopback0
  ip address 20.1.1.2/24
  ip pim sparse-mode
  ip igmp static-oif 225.1.1.1

hardware access-list tcam region mcast-nat 512

<<Ingress NAT>>

ip route 30.1.1.0/24 10.1.1.1
ip pim ssm range 232.0.0.0/8
ip service-reflect source-interface loopback0
ip service-reflect mode ingress 235.1.1.0/24
ip service-reflect destination 235.1.1.1 to 234.1.1.1 mask-len 32 source 30.1.1.70 to
20.1.1.70 mask-len 32
hardware access-list tcam region mcast-nat 512

<<Egress NAT>>

ip route 30.1.1.0/24 10.1.1.1
ip pim ssm range 232.0.0.0/8
ip service-reflect mode egress 225.1.1.0/24
ip service-reflect destination 225.1.1.1 to 224.1.1.1 mask-len 32 source 30.1.1.1 to 20.1.1.1
 mask-len 32 static-oif port-channel40
ip service-reflect destination 225.1.1.1 to 224.1.1.100 mask-len 32 source 30.1.1.1 to
20.1.1.100 mask-len 32 static-oif port-channel40
ip service-reflect destination 225.1.1.1 to 224.1.1.101 mask-len 32 source 30.1.1.1 to
20.1.1.101 mask-len 32 static-oif port-channel40
ip service-reflect destination 235.1.1.1 to 234.1.1.1 mask-len 32 source 30.1.1.70 to
20.1.1.70 mask-len 32
multicast service-reflect interface all map interface Ethernet1/21
hardware access-list tcam region mcast-nat 512
interface Ethernet1/21
  link loopback
  no shutdown
```

```
interface Ethernet1/21.1
  encapsulation dot1q 10
  no shutdown
interface Ethernet1/21.2
  encapsulation dot1q 20
  no shutdown
interface Ethernet1/21.3
  encapsulation dot1q 30
  no shutdown
interface Ethernet1/21.4
  encapsulation dot1q 40
  no shutdown
```

The following examples show the display/output of the Multicast Service Reflection show commands:

```
switch# show ip mroute sr
IP Multicast Routing Table for VRF "default"
(30.1.1.1/32, 225.1.1.1/32), uptime: 01:29:45, ip mrib pim
  NAT Mode: Egress
  NAT Route Type: Pre
  Incoming interface: Ethernet1/1, RPF nbr: 10.1.1.1
  Outgoing interface list: (count: 1)
    loopback0, uptime: 01:29:45, mrib
       SR: (20.1.1.1, 224.1.1.1) OIF: port-channel40
       SR: (20.1.1.100, 224.1.1.100) OIF: port-channel40
       SR: (20.1.1.101, 224.1.1.101) OIF: port-channel40

(30.1.1.70/32, 235.1.1.1/32), uptime: 01:05:12, ip mrib pim
  NAT Mode: Ingress
  NAT Route Type: Pre
  Incoming interface: Ethernet1/1, RPF nbr: 10.1.1.1
  Outgoing interface list: (count: 1)
    loopback0, uptime: 01:05:12, mrib
       SR: (20.1.1.70, 234.1.1.1)
```

```
switch# show ip mroute 234.1.1.1 detail
IP Multicast Routing Table for VRF "default"
Total number of routes: 26
Total number of (*,G) routes: 19
Total number of (S,G) routes: 6
Total number of (*,G-prefix) routes: 1

(20.1.1.70/32, 234.1.1.1/32), uptime: 01:06:30, mrib(0) ip(0) pim(0) static(1)
  RPF-Source: 20.1.1.70 [0/0]
  Data Created: Yes
  Stats: 499/24259 [Packets/Bytes], 27.200  bps
  Stats: Active Flow
  Incoming interface: loopback0, RPF nbr: 20.1.1.70
  LISP dest context id: 0  Outgoing interface list: (count: 1) (bridge-only: 0)
    port-channel40, uptime: 00:59:20, static
```

```
switch# show forwarding distribution multicast route
IPv4 Multicast Routing Table for table-id: 1
Total number of groups: 22
Legend:
   C = Control Route
   D = Drop Route
   G = Local Group (directly connected receivers)
   O = Drop on RPF Fail
   P = Punt to supervisor
   L = SRC behind L3
   d = Decap Route
   Es = Extranet src entry
   Er = Extranet recv entry
   Nf = VPC None-Forwarder
```

```
 dm = MVPN Decap Route
 em = MVPN Encap Route
 IPre = Ingress Service-reflect Pre
 EPre = Egress Service-reflect Pre
 Pst = Ingress/Egress Service-reflect Post

(30.1.1.70/32, 235.1.1.1/32), RPF Interface: Ethernet1/1, flags: IPre
  Upstream Nbr: 10.1.1.1
  Received Packets: 25 Bytes: 1625
  Number of Outgoing Interfaces: 1
  Outgoing Interface List Index: 4
    port-channel40

(20.1.1.1/32, 224.1.1.1/32), RPF Interface: loopback0, flags: Pst
  Upstream Nbr: 20.1.1.1
  Received Packets: 0 Bytes: 0
  Number of Outgoing Interfaces: 1
  Outgoing Interface List Index: 2
    port-channel40

(20.1.1.100/32, 224.1.1.100/32), RPF Interface: loopback0, flags: Pst
  Upstream Nbr: 20.1.1.100
  Received Packets: 0 Bytes: 0
  Number of Outgoing Interfaces: 1
  Outgoing Interface List Index: 2
    port-channel40

(20.1.1.101/32, 224.1.1.101/32), RPF Interface: loopback0, flags: Pst
  Upstream Nbr: 20.1.1.101
  Received Packets: 0 Bytes: 0
  Number of Outgoing Interfaces: 1
  Outgoing Interface List Index: 2
    port-channel40
```

```
switch# show forwarding multicast route group 235.1.1.1 source 30.1.1.70
slot  1
=======
  (30.1.1.70/32, 235.1.1.1/32), RPF Interface: Ethernet1/1, flags: c
    Received Packets: 18 Bytes: 1170
    Outgoing Interface List Index: 4
    Number of next hops: 1
    oiflist flags: 16384
  Outgoing Interface List Index: 0x4
    port-channel40
```

Configuration Examples for PIM

This section describes how to configure PIM using different data distribution modes and RP selection methods.

SSM Configuration Example

To configure PIM in SSM mode, follow these steps for each router in the PIM domain:

1. Configure PIM sparse mode parameters on the interfaces that you want to participate in the domain. We recommend that you enable PIM on all interfaces.

```
switch# configure terminal
switch(config)# interface ethernet 2/1
```

```
switch(config-if)# ip pim sparse-mode
```

2. Configure the parameters for IGMP that support SSM. Usually, you configure IGMPv3 on PIM interfaces to support SSM.

```
switch# configure terminal
switch(config)# interface ethernet 2/1
switch(config-if)# ip igmp version 3
```

3. Configure the SSM range if you do not want to use the default range.

```
switch# configure terminal
switch(config)# ip pim ssm range 239.128.1.0/24
```

4. Configure message filtering.

```
switch# configure terminal
switch(config)# ip pim log-neighbor-changes
```

The following example shows how to configure PIM SSM mode:

```
configure terminal
  interface ethernet 2/1
    ip pim sparse-mode
    ip igmp version 3
    exit
  ip pim ssm range 239.128.1.0/24
  ip pim log-neighbor-changes
```

PIM SSM Over vPC Configuration Example

This example shows how to override the default SSM range of 232.0.0.0/8 to 225.1.1.0/24. PIM SSM over vPC will work as long as S,G joins are received in this range.

```
switch# configure terminal
switch(config)# vrf context Enterprise
switch(config-vrf)# ip pim ssm range 225.1.1.0/24
switch(config-vrf)# show ip pim group-range --> Shows the configured SSM group range.
PIM Group-Range Configuration for VRF "Enterprise"
Group-range        Mode      RP-address       Shared-tree-only range
225.1.1.0/24       SSM       -                -

switch1# show vpc (primary vPC) --> Shows vPC-related information.
Legend:
                (*) - local vPC is down, forwarding via vPC peer-link

vPC domain id                     : 10
Peer status                       : peer adjacency formed ok
vPC keep-alive status             : peer is alive
Configuration consistency status  : success
Per-vlan consistency status       : success
Type-2 consistency status         : success
vPC role                          : primary
Number of vPCs configured         : 2
Peer Gateway                      : Disabled
```

```
Dual-active excluded VLANs        : -
Graceful Consistency Check        : Enabled
Auto-recovery status              : Disabled
Delay-restore status              : Timer is off.(timeout = 30s)
Delay-restore SVI status          : Timer is off.(timeout = 10s)

vPC Peer-link status
---------------------------------------------------------------------
id   Port   Status Active vlans
--   ----   ------ --------------------------------------------------
1    Po1000 up     101-102

vPC status
----------------------------------------------------------------------------
id   Port   Status Consistency Reason                     Active vlans
--   ----   ------ ----------- ------                     ------------
1    Po1    up     success     success                    102

2    Po2    up     success     success                    101

switch2# show vpc (secondary vPC)
Legend:
                (*) - local vPC is down, forwarding via vPC peer-link

vPC domain id                     : 10
Peer status                       : peer adjacency formed ok
vPC keep-alive status             : peer is alive
Configuration consistency status  : success
Per-vlan consistency status       : success
Type-2 consistency status         : success
vPC role                          : secondary
Number of vPCs configured         : 2
Peer Gateway                      : Disabled
Dual-active excluded VLANs        : -
Graceful Consistency Check        : Enabled
Auto-recovery status              : Disabled
Delay-restore status              : Timer is off.(timeout = 30s)
Delay-restore SVI status          : Timer is off.(timeout = 10s)

vPC Peer-link status
---------------------------------------------------------------------
id   Port   Status Active vlans
--   ----   ------ --------------------------------------------------
1    Po1000 up     101-102

vPC status
----------------------------------------------------------------------------
id   Port   Status Consistency Reason                     Active vlans
--   ----   ------ ----------- ------                     ------------
1    Po1    up     success     success                    102

2    Po2    up     success     success                    101


switch1# show ip igmp snooping group vlan 101 (primary vPC IGMP snooping states) --> Shows
 if S,G v3 joins are received and on which VLAN. The same VLAN should be OIF in the MRIB
output.

Type: S - Static, D - Dynamic, R - Router port, F - Fabricpath core port

Vlan  Group Address      Ver  Type  Port list
101   */*                -    R     Po1000 Vlan101
101   225.1.1.1          v3
        100.6.160.20           D     Po2
```

```
switch2# show ip igmp snooping group vlan 101 (secondary vPC IGMP snooping states)
Type: S - Static, D - Dynamic, R - Router port, F - Fabricpath core port

Vlan  Group Address      Ver  Type  Port list
101   */*                -    R     Po1000 Vlan101
101   225.1.1.1          v3
        100.6.160.20           D     Po2


switch1# show ip pim route (primary vPC PIM route) --> Shows the route information in the
PIM protocol.
PIM Routing Table for VRF "default" - 3 entries

(10.6.159.20/32, 225.1.1.1/32), expires 00:02:37
  Incoming interface: Ethernet1/19, RPF nbr 10.6.159.20
  Oif-list:       (1) 00000000, timeout-list: (0) 00000000
  Immediate-list: (1) 00000000, timeout-list: (0) 00000000
  Sgr-prune-list: (0) 00000000
  Timeout-interval: 2, JP-holdtime round-up: 3

(100.6.160.20/32, 225.1.1.1/32), expires 00:01:19
  Incoming interface: Vlan102, RPF nbr 100.6.160.20
  Oif-list:       (0) 00000000, timeout-list: (0) 00000000
  Immediate-list: (0) 00000000, timeout-list: (0) 00000000
  Sgr-prune-list: (0) 00000000
  Timeout-interval: 2, JP-holdtime round-up: 3

(*, 232.0.0.0/8), expires 00:01:19
  Incoming interface: Null0, RPF nbr 0.0.0.0
  Oif-list:       (0) 00000000, timeout-list: (0) 00000000
  Immediate-list: (0) 00000000, timeout-list: (0) 00000000
  Sgr-prune-list: (0) 00000000
  Timeout-interval: 2, JP-holdtime round-up: 3

switch2# show ip pim route (secondary vPC PIM route)
PIM Routing Table for VRF "default" - 3 entries
(10.6.159.20/32, 225.1.1.1/32), expires 00:02:51
  Incoming interface: Vlan102, RPF nbr 100.6.160.100
  Oif-list:       (0) 00000000, timeout-list: (0) 00000000
  Immediate-list: (0) 00000000, timeout-list: (0) 00000000
  Sgr-prune-list: (0) 00000000
  Timeout-interval: 3, JP-holdtime round-up: 3

(100.6.160.20/32, 225.1.1.1/32), expires 00:02:51
  Incoming interface: Vlan102, RPF nbr 100.6.160.20
  Oif-list:       (0) 00000000, timeout-list: (0) 00000000
  Immediate-list: (0) 00000000, timeout-list: (0) 00000000
  Sgr-prune-list: (0) 00000000
  Timeout-interval: 3, JP-holdtime round-up: 3

(*, 232.0.0.0/8), expires 00:02:51
  Incoming interface: Null0, RPF nbr 0.0.0.0
  Oif-list:       (0) 00000000, timeout-list: (0) 00000000
  Immediate-list: (0) 00000000, timeout-list: (0) 00000000
  Sgr-prune-list: (0) 00000000
  Timeout-interval: 3, JP-holdtime round-up: 3

switch2# show ip pim route (secondary vPC PIM route)
PIM Routing Table for VRF "default" - 3 entries

(10.6.159.20/32, 225.1.1.1/32), expires 00:02:29
  Incoming interface: Vlan102, RPF nbr 100.6.160.100
  Oif-list:       (0) 00000000, timeout-list: (0) 00000000
```

```
  Immediate-list: (0) 00000000, timeout-list: (0) 00000000
  Sgr-prune-list: (0) 00000000
  Timeout-interval: 3, JP-holdtime round-up: 3

(100.6.160.20/32, 225.1.1.1/32), expires 00:02:29
  Incoming interface: Vlan102, RPF nbr 100.6.160.20
  Oif-list:       (0) 00000000, timeout-list: (0) 00000000
  Immediate-list: (0) 00000000, timeout-list: (0) 00000000
  Sgr-prune-list: (0) 00000000
  Timeout-interval: 3, JP-holdtime round-up: 3

(*, 232.0.0.0/8), expires 00:02:29
  Incoming interface: Null0, RPF nbr 0.0.0.0
  Oif-list:       (0) 00000000, timeout-list: (0) 00000000
  Immediate-list: (0) 00000000, timeout-list: (0) 00000000
  Sgr-prune-list: (0) 00000000
  Timeout-interval: 3, JP-holdtime round-up: 3


switch1# show ip mroute (primary vPC MRIB route) --> Shows the IP multicast routing table.

IP Multicast Routing Table for VRF "default"

(10.6.159.20/32, 225.1.1.1/32), uptime: 03:16:40, pim ip
  Incoming interface: Ethernet1/19, RPF nbr: 10.6.159.20
  Outgoing interface list: (count: 1)
    Vlan102, uptime: 03:16:40, pim

(100.6.160.20/32, 225.1.1.1/32), uptime: 03:48:57, igmp ip pim
  Incoming interface: Vlan102, RPF nbr: 100.6.160.20
  Outgoing interface list: (count: 1)
    Vlan101, uptime: 03:48:57, igmp

(*, 232.0.0.0/8), uptime: 6d06h, pim ip
  Incoming interface: Null, RPF nbr: 0.0.0.0
  Outgoing interface list: (count: 0)

switch1# show ip mroute detail (primary vPC MRIB route) --> Shows if the (S,G) entries have
 the RPF as the interface toward the source and no *,G states are maintained for the SSM
group range in the MRIB.

IP Multicast Routing Table for VRF "default"

Total number of routes: 3
Total number of (*,G) routes: 0
Total number of (S,G) routes: 2
Total number of (*,G-prefix) routes: 1

(10.6.159.20/32, 225.1.1.1/32), uptime: 03:24:28, pim(1) ip(0)
  Data Created: Yes
  VPC Flags
    RPF-Source Forwarder
  Stats: 1/51 [Packets/Bytes], 0.000   bps
  Stats: Inactive Flow
  Incoming interface: Ethernet1/19, RPF nbr: 10.6.159.20
  Outgoing interface list: (count: 1)
    Vlan102, uptime: 03:24:28, pim

(100.6.160.20/32, 225.1.1.1/32), uptime: 03:56:45, igmp(1) ip(0) pim(0)
  Data Created: Yes
  VPC Flags
    RPF-Source Forwarder
  Stats: 1/51 [Packets/Bytes], 0.000   bps
  Stats: Inactive Flow
```

```
  Incoming interface: Vlan102, RPF nbr: 100.6.160.20
  Outgoing interface list: (count: 1)
    Vlan101, uptime: 03:56:45, igmp (vpc-svi)

(*, 232.0.0.0/8), uptime: 6d06h, pim(0) ip(0)
  Data Created: No
  Stats: 0/0 [Packets/Bytes], 0.000   bps
  Stats: Inactive Flow
  Incoming interface: Null, RPF nbr: 0.0.0.0
  Outgoing interface list: (count: 0)

switch2# show ip mroute detail (secondary vPC MRIB route)
IP Multicast Routing Table for VRF "default"

Total number of routes: 3
Total number of (*,G) routes: 0
Total number of (S,G) routes: 2
Total number of (*,G-prefix) routes: 1

(10.6.159.20/32, 225.1.1.1/32), uptime: 03:26:24, igmp(1) pim(0) ip(0)
  Data Created: Yes
  Stats: 1/51 [Packets/Bytes], 0.000   bps
  Stats: Inactive Flow
  Incoming interface: Vlan102, RPF nbr: 100.6.160.100
  Outgoing interface list: (count: 1)
    Ethernet1/17, uptime: 03:26:24, igmp

(100.6.160.20/32, 225.1.1.1/32), uptime: 04:06:32, igmp(1) ip(0) pim(0)
  Data Created: Yes
  VPC Flags
    RPF-Source Forwarder
  Stats: 1/51 [Packets/Bytes], 0.000   bps
  Stats: Inactive Flow
  Incoming interface: Vlan102, RPF nbr: 100.6.160.20
  Outgoing interface list: (count: 1)
    Vlan101, uptime: 04:03:24, igmp (vpc-svi)

(*, 232.0.0.0/8), uptime: 6d06h, pim(0) ip(0)
  Data Created: No
  Stats: 0/0 [Packets/Bytes], 0.000   bps
  Stats: Inactive Flow
  Incoming interface: Null, RPF nbr: 0.0.0.0
  Outgoing interface list: (count: 0)
```

BSR Configuration Example

To configure PIM in ASM mode using the BSR mechanism, follow these steps for each router in the PIM domain:

1. Configure PIM sparse mode parameters on the interfaces that you want to participate in the domain. We recommend that you enable PIM on all interfaces.

```
switch# configure terminal
switch(config)# interface ethernet 2/1
switch(config-if)# ip pim sparse-mode
```

2. Configure whether that router should listen and forward BSR messages.

```
switch# configure terminal
```

```
switch(config)# ip pim bsr forward listen
```

3. Configure the BSR parameters for each router that you want to act as a BSR.

```
switch# configure terminal
switch(config)# ip pim bsr-candidate ethernet 2/1 hash-len 30
```

4. Configure the RP parameters for each router that you want to act as a candidate RP.

```
switch# configure terminal
switch(config)# ip pim rp-candidate ethernet 2/1 group-list 239.0.0.0/24
```

5. Configure message filtering.

```
switch# configure terminal
switch(config)# ip pim log-neighbor-changes
```

The following example shows how to configure PIM ASM mode using the BSR mechanism and how to configure the BSR and RP on the same router:

```
configure terminal
  interface ethernet 2/1
    ip pim sparse-mode
    exit
  ip pim bsr forward listen
ip pim bsr-candidate ethernet 2/1 hash-len 30
  ip pim rp-candidate ethernet 2/1 group-list 239.0.0.0/24
  ip pim log-neighbor-changes
```

Auto-RP Configuration Example

To configure PIM in Bidir mode using the Auto-RP mechanism, follow these steps for each router in the PIM domain:

1. Configure PIM sparse mode parameters on the interfaces that you want to participate in the domain. We recommend that you enable PIM on all interfaces.

```
switch# configure terminal
switch(config)# interface ethernet 2/1
switch(config-if)# ip pim sparse-mode
```

2. Configure whether that router should listen and forward Auto-RP messages.

```
switch# configure terminal
switch(config)# ip pim auto-rp forward listen
```

3. Configure the mapping agent parameters for each router that you want to act as a mapping agent.

```
switch# configure terminal
switch(config)# ip pim auto-rp mapping-agent ethernet 2/1
```

4. Configure the RP parameters for each router that you want to act as a candidate RP.

```
switch# configure terminal
switch(config)# ip pim auto-rp rp-candidate ethernet 2/1 group-list 239.0.0.0/24 bidir
```

5. Configure message filtering.

```
switch# configure terminal
switch(config)# ip pim log-neighbor-changes
```

This example shows how to configure PIM Bidir mode using the Auto-RP mechanism and how to configure the mapping agent and RP on the same router:

```
configure terminal
  interface ethernet 2/1
    ip pim sparse-mode
    exit
  ip pim auto-rp listen
  ip pim auto-rp forward
  ip pim auto-rp mapping-agent ethernet 2/1
  ip pim auto-rp rp-candidate ethernet 2/1 group-list 239.0.0.0/24 bidir
  ip pim log-neighbor-changes
```

PIM Anycast RP Configuration Example

To configure ASM mode using the PIM Anycast-RP method, follow these steps for each router in the PIM domain:

1. Configure PIM sparse mode parameters on the interfaces that you want to participate in the domain. We recommend that you enable PIM on all interfaces.

```
switch# configure terminal
switch(config)# interface ethernet 2/1
switch(config-if)# ip pim sparse-mode
```

2. Configure the RP address that you configure on all routers in the Anycast-RP set.

```
switch# configure terminal
switch(config)# interface loopback 0
switch(config-if)# ip address 192.0.2.3/32
switch(config-if)# ip pim sparse-mode
```

3. Configure a loopback with an address to use in communication between routers in the Anycast-RP set for each router that you want to be in the Anycast-RP set.

```
switch# configure terminal
switch(config)# interface loopback 1
switch(config-if)# ip address 192.0.2.31/32
switch(config-if)# ip pim sparse-mode
```

4. Configure the Anycast-RP parameters and repeat with the IP address of each Anycast-RP for each router that you want to be in the Anycast-RP set. This example shows two Anycast-RPs.

```
switch# configure terminal
switch(config)# ip pim anycast-rp 192.0.2.3 193.0.2.31
switch(config)# ip pim anycast-rp 192.0.2.3 193.0.2.32
```

5. Configure message filtering.

```
switch# configure terminal
switch(config)# ip pim log-neighbor-changes
```

The following example shows how to configure PIM Anycast RP for IPv6:

```
configure terminal
interface loopback 0
ipv6 address 2001:0db8:0:abcd::5/32
ipv6 pim sparse-mode
ipv6 router ospfv3 1 area 0.0.0.0
exit
interface loopback 1
ipv6 address 2001:0db8:0:abcd::1111/32
ipv6 pim sparse-mode
ipv6 router ospfv3 1 area 0.0.0.0
exit
ipv6 pim rp-address 2001:0db8:0:abcd::1111 group-list ff1e:abcd:def1::0/24
ipv6 pim anycast-rp 2001:0db8:0:abcd::5 2001:0db8:0:abcd::1111
```

The following example shows how to configure PIM ASM mode using two Anycast-RPs:

```
configure terminal
interface ethernet 2/1
ip pim sparse-mode
exit
interface loopback 0
ip address 192.0.2.3/32
ip pim sparse-mode
exit
interface loopback 1
ip address 192.0.2.31/32
ip pim sparse-mode
exit
ip pim anycast-rp 192.0.2.3 192.0.2.31
ip pim anycast-rp 192.0.2.3 192.0.2.32
ip pim log-neighbor-changes
```

Prefix-Based and Route-Map-Based Configurations

```
ip prefix-list plist11 seq 10 deny 231.129.128.0/17
ip prefix-list plist11 seq 20 deny 231.129.0.0/16
ip prefix-list plist11 seq 30 deny 231.128.0.0/9
ip prefix-list plist11 seq 40 permit 231.0.0.0/8

ip prefix-list plist22 seq 10 deny 231.129.128.0/17
ip prefix-list plist22 seq 20 deny 231.129.0.0/16
ip prefix-list plist22 seq 30 permit 231.128.0.0/9
ip prefix-list plist22 seq 40 deny 231.0.0.0/8

ip prefix-list plist33 seq 10 deny 231.129.128.0/17
ip prefix-list plist33 seq 20 permit 231.129.0.0/16
ip prefix-list plist33 seq 30 deny 231.128.0.0/9
ip prefix-list plist33 seq 40 deny 231.0.0.0/8

ip pim rp-address 172.21.0.11 prefix-list plist11
ip pim rp-address 172.21.0.22 prefix-list plist22
ip pim rp-address 172.21.0.33 prefix-list plist33
route-map rmap11 deny 10
 match ip multicast group 231.129.128.0/17
route-map rmap11 deny 20
```

```
 match ip multicast group 231.129.0.0/16
route-map rmap11 deny 30
 match ip multicast group 231.128.0.0/9
route-map rmap11 permit 40
 match ip multicast group 231.0.0.0/8

route-map rmap22 deny 10
 match ip multicast group 231.129.128.0/17
route-map rmap22 deny 20
 match ip multicast group 231.129.0.0/16
route-map rmap22 permit 30
 match ip multicast group 231.128.0.0/9
route-map rmap22 deny 40
 match ip multicast group 231.0.0.0/8

route-map rmap33 deny 10
 match ip multicast group 231.129.128.0/17
route-map rmap33 permit 20
 match ip multicast group 231.129.0.0/16
route-map rmap33 deny 30
 match ip multicast group 231.128.0.0/9
route-map rmap33 deny 40
 match ip multicast group 231.0.0.0/8

ip pim rp-address 172.21.0.11 route-map rmap11
ip pim rp-address 172.21.0.22 route-map rmap22
ip pim rp-address 172.21.0.33 route-map rmap33
```

Output

```
dc3rtg-d2(config-if)# show ip pim rp
PIM RP Status Information for VRF "default"
BSR disabled
Auto-RP disabled
BSR RP Candidate policy: None
BSR RP policy: None
Auto-RP Announce policy: None
Auto-RP Discovery policy: None

RP: 172.21.0.11, (0), uptime: 00:12:36, expires: never,
  priority: 0, RP-source: (local), group-map: rmap11, group ranges:
      231.0.0.0/8  231.128.0.0/9 (deny)
      231.129.0.0/16 (deny) 231.129.128.0/17 (deny)
RP: 172.21.0.22, (0), uptime: 00:12:36, expires: never,
  priority: 0, RP-source: (local), group-map: rmap22, group ranges:
      231.0.0.0/8 (deny) 231.128.0.0/9
      231.129.0.0/16 (deny) 231.129.128.0/17 (deny)
RP: 172.21.0.33, (0), uptime: 00:12:36, expires: never,
  priority: 0, RP-source: (local), group-map: rmap33, group ranges:
      231.0.0.0/8 (deny) 231.128.0.0/9 (deny)
      231.129.0.0/16  231.129.128.0/17 (deny)

dc3rtg-d2(config-if)# show ip mroute
IP Multicast Routing Table for VRF "default"

(*, 231.1.1.1/32), uptime: 00:07:20, igmp pim ip
  Incoming interface: Ethernet2/1, RPF nbr: 10.165.20.1
  Outgoing interface list: (count: 1)
    loopback1, uptime: 00:07:20, igmp

(*, 231.128.1.1/32), uptime: 00:14:27, igmp pim ip
  Incoming interface: Ethernet2/1, RPF nbr: 10.165.20.1
  Outgoing interface list: (count: 1)
```

```
    loopback1, uptime: 00:14:27, igmp

(*, 231.129.1.1/32), uptime: 00:14:25, igmp pim ip
  Incoming interface: Ethernet2/1, RPF nbr: 10.165.20.1
  Outgoing interface list: (count: 1)
    loopback1, uptime: 00:14:25, igmp

(*, 231.129.128.1/32), uptime: 00:14:26, igmp pim ip
  Incoming interface: Null, RPF nbr: 10.0.0.1
  Outgoing interface list: (count: 1)
    loopback1, uptime: 00:14:26, igmp

(*, 232.0.0.0/8), uptime: 1d20h, pim ip
  Incoming interface: Null, RPF nbr: 10.0.0.1
  Outgoing interface list: (count: 0)

dc3rtg-d2(config-if)# show ip pim group-range
PIM Group-Range Configuration for VRF "default"
Group-range        Mode      RP-address       Shared-tree-only range
232.0.0.0/8        ASM       -                -
231.0.0.0/8        ASM       172.21.0.11      -
231.128.0.0/9      ASM       172.21.0.22      -
231.129.0.0/16     ASM       172.21.0.33      -
231.129.128.0/17   Unknown   -                -
```

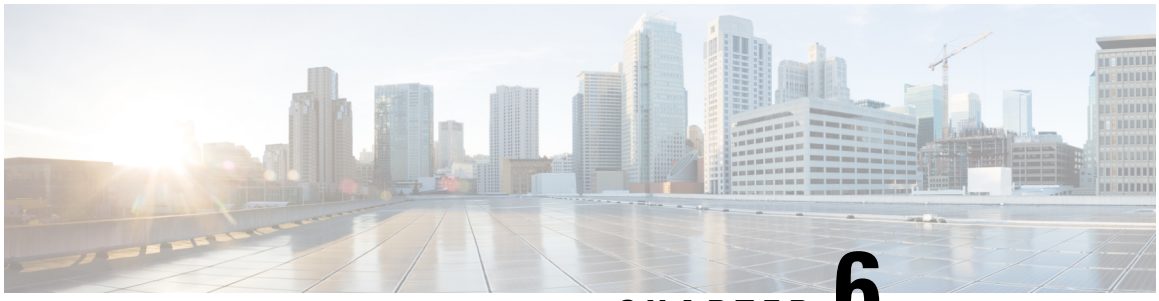
Related Documents

Related Topic	Document Title
Configuring VRFs	*Cisco Nexus 9000 Series NX-OS Unicast Routing Configu*

Standards

MIBs

MIBs	MIBs Link
MIBs related to PIM	To locate and download supported MIBs, go to the follow https://cisco.github.io/cisco-mibs/supportlists/nexus9000/Nexus9000MIBSupportList.html



CHAPTER 6

Configuring IGMP Snooping

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

- About IGMP Snooping, on page 129
- Prerequisites for IGMP Snooping, on page 131
- Guidelines and Limitations for IGMP Snooping, on page 132
- Default Settings, on page 133
- Configuring IGMP Snooping Parameters, on page 133
- Verifying the IGMP Snooping Configuration, on page 140
- Displaying IGMP Snooping Statistics, on page 140
- Clearing IGMP Snooping Statistics, on page 140
- Configuration Examples for IGMP Snooping, on page 140

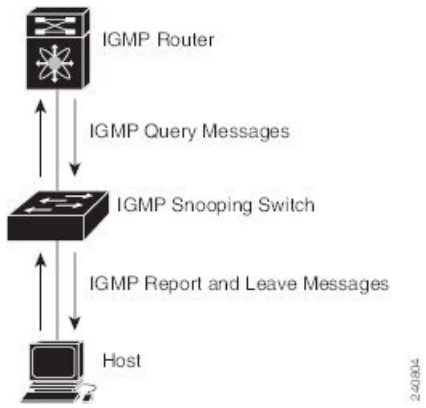
About IGMP Snooping

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 Cisco 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 Cisco 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 *Cisco Nexus 9000 Series NX-OS Unicast Routing Configuration 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:

- Cisco Nexus 9000 Series 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.
- Cisco Nexus 9508 and 9504 platform switches with N9K-X9636C-R, N9K-X9636Q-R, and N9K-X9636C-RX line cards support IGMP snooping with vPCs.
- 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 Cisco NX-OS Release 7.0(3)I3(1) use the **mvr-suppress-query vlan** *<id>* command.

- In releases prior to Cisco 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
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 Cisco IOS CLI, be aware that the Cisco NX-OS commands for this feature might differ from the Cisco 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 Cisco 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 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 `switch(config-vlan-config)# ip igmp snooping fast-leave`	Supports IGMPv2 hosts that cannot be explicitly tracked because of the host report suppression mechanism of the 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.
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.

Option	Description
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* `switch(config-vlan-config)# ip igmp snooping query-interval 120`	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.
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.
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.
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.

Option	Description
`switch(config-vlan-config)# ip igmp snooping report-suppression`	
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.

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 *Cisco Nexus 9000 Series 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
```

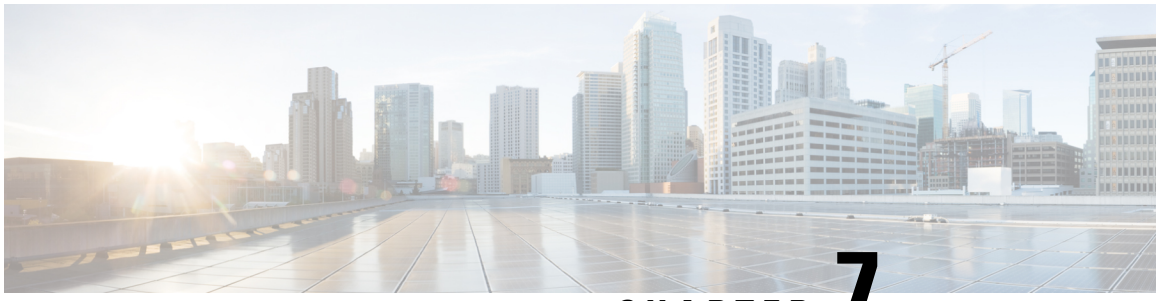
In the above example, the prefix-list permits 224.1.1.1 and 224.1.1.2 but rejects 224.1.1.3 and all the groups in the 225.0.0.0/8 range. The prefix-list is an implicit "deny" if there is no match. If you wish to permit everything else, add **ip prefix-list plist seq 30 permit 224.0.0.0/4 eq 32**.

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 7

Configuring MSDP

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

- About MSDP, on page 143
- Prerequisites for MSDP, on page 145
- Default Settings, on page 145
- Configuring MSDP, on page 146
- Verifying the MSDP Configuration, on page 154
- Monitoring MSDP, on page 155
- Configuration Examples for MSDP, on page 156
- Related Documents, on page 157
- Standards, on page 157

About MSDP

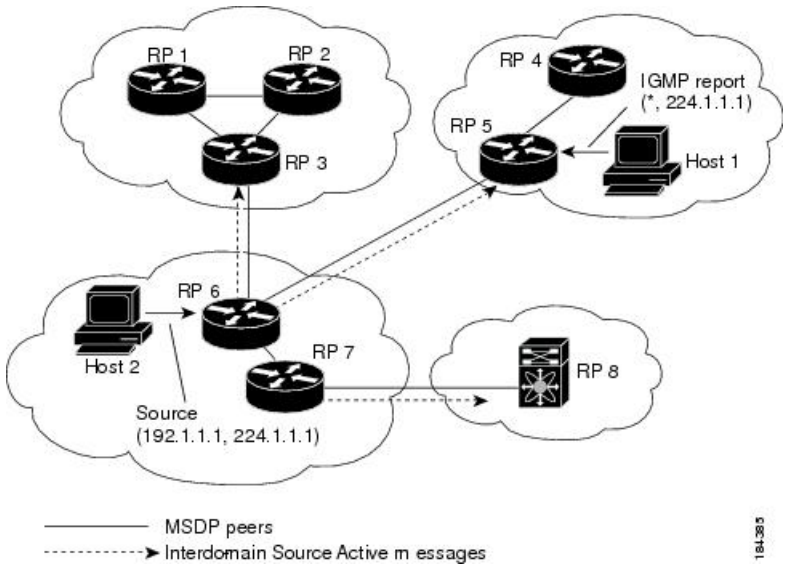
You can use the Multicast Source Discovery Protocol (MSDP) to exchange multicast source information between multiple Border Gateway Protocol (BGP) enabled Protocol Independent Multicast (PIM) sparse-mode domains. In addition, MSDP can be used to create an Anycast-RP configuration to provide RP redundancy and load sharing. For information about BGP, see the *Cisco Nexus 9000 Series NX-OS Unicast Routing Configuration Guide*.

MSDP is supported on all Cisco Nexus 9000 series switches.

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.

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.

Default Settings

This table lists the default settings for MSDP parameters.

Table 18: 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

Parameters	Default
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 Cisco IOS CLI, be aware that the Cisco NX-OS commands for this feature might differ from the Cisco IOS commands that you would use.

Enabling the MSDP Feature

SUMMARY STEPS

1. **configure terminal**
2. **feature msdp**
3. (Optional) **show running-configuration msdp**
4. (Optional) **copy running-config startup-config**

DETAILED STEPS

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.

SUMMARY STEPS

1. **configure terminal**
2. **ip msdp peer** *peer-ip-address* **connect-source** *interface* [**remote-as** *as-number*]
3. Repeat Step 2 for each MSDP peering relationship by changing the peer IP address, the interface, and the AS number as appropriate.
4. (Optional) **show ip msdp summary** [**vrf** [*vrf-name* | **all**]]
5. (Optional) **copy running-config startup-config**

DETAILED STEPS

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 MSDP 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 19: 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 *Cisco Nexus 9000 Series 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 *Cisco Nexus 9000 Series 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.

SUMMARY STEPS

1. **configure terminal**
2. **ip msdp description** *peer-ip-address description*
3. **ip msdp shutdown** *peer-ip-address*
4. **ip msdp password** *peer-ip-address password*
5. **ip msdp sa-policy** *peer-ip-address policy-name* **in**
6. **ip msdp sa-policy** *peer-ip-address policy-name* **out**
7. **ip msdp sa-limit** *peer-ip-address limit*
8. (Optional) **show ip msdp peer** [*peer-address*] [**vrf** [*vrf-name* | **all**]]
9. (Optional) **copy running-config startup-config**

DETAILED STEPS

Procedure

	Command or Action	Purpose

Step 1	**configure terminal**
	Example:
	`switch# configure terminal`
	`switch(config)#`
		Enters global configuration mode.
		Note
		Use the commands listed from step-2 to configure the MSDP peer parameters.

Step 2	**ip msdp description** *peer-ip-address description* **Example:** `switch(config)# ip msdp description 192.168.1.10 peer in Engineering network`	Sets a description string for the peer. By default, the peer has no description.
Step 3	**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.
Step 4	**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.
Step 5	**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.
Step 6	**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.
Step 7	**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.
Step 8	(Optional) **show ip msdp peer** [*peer-address*] [**vrf** [*vrf-name* | **all**]] **Example:** `switch(config)# show ip msdp peer 192.168.1.10`	Displays detailed MSDP peer information.
Step 9	(Optional) **copy running-config startup-config** **Example:**	Copies the running configuration to the startup configuration.

	Command or Action	Purpose
	`switch(config)# copy running-config startup-config`	

Configuring MSDP Global Parameters

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

Table 20: 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.

SUMMARY STEPS

1. **configure terminal**
2. **ip msdp originator-id** *interface*
3. **ip msdp group-limit** *limit* **source** *source-prefix*
4. **ip msdp sa-interval** *seconds*
5. (Optional) **show ip msdp summary** [**vrf** [*vrf-name* | **all**]]
6. (Optional) **copy running-config startup-config**

DETAILED STEPS

Procedure

	Command or Action	Purpose

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

Step 2	**ip msdp originator-id** *interface* **Example:** `switch(config)# ip msdp originator-id loopback0`	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.
Step 3	**ip msdp group-limit** *limit* **source** *source-prefix* **Example:** `switch(config)# ip msdp group-limit 1000 source 192.168.1.0/24`	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.
Step 4	**ip msdp sa-interval** *seconds* **Example:** `switch(config)# ip msdp sa-interval 80`	Interval at which the software transmits Source-Active (SA) messages. The range is from 60 to 65,535 seconds. The default is 60 seconds.
Step 5	(Optional) **show ip msdp summary** [**vrf** [*vrf-name* | **all**]] **Example:** `switch(config)# show ip msdp summary`	Displays a summary of the MSDP configuration.
Step 6	(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 MSDP 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.

SUMMARY STEPS

1. **configure terminal**
2. **ip msdp mesh-group** *peer-ip-addr mesh-name*
3. Repeat Step 2 for each MSDP peer in the mesh by changing the peer IP address.
4. (Optional) **show ip msdp mesh-group** [*mesh-group*] [**vrf** [*vrf-name* | **all**]]
5. (Optional) **copy running-config startup-config**

DETAILED STEPS

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 MSDP 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.

SUMMARY STEPS

1. **restart msdp**
2. **configure terminal**
3. **ip msdp flush-routes**
4. (Optional) **show running-configuration | include flush-routes**

5. (Optional) **copy running-config startup-config**

DETAILED STEPS

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.
show ip msdp peer [*peer-address*] [**vrf** [*vrf-name* | **all**]]	Displays MSDP information for the MSDP peer.

Command	Description
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)

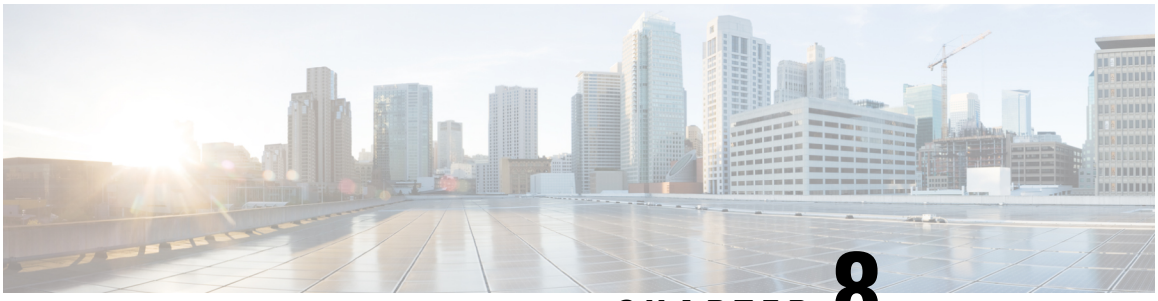
```
configure terminal
  ip msdp peer 192.168.7.10 connect-source ethernet 1/1
  ip msdp peer 192.168.3.10 connect-source ethernet 1/2 remote-as
7
  ip msdp peer 192.168.5.10 connect-source ethernet 1/3 remote-as
8
  ip msdp password 192.168.3.10 my_peer_password_36
  ip msdp password 192.168.5.10 my_peer_password_56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

Related Documents

Related Topic	Document Title
Configuring MBGP	*Cisco Nexus 9000 Series NX-OS Unicast Routing Configu*

Standards

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



CHAPTER 8

Configuring MVR

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

This chapter contains the following sections:

- About MVR, on page 159
- MVR Interoperation with Other Features, on page 160
- Guidelines and Limitations for MVR, on page 160
- Default MVR Settings, on page 160
- Configuring MVR, on page 161
- Verifying the MVR Configuration, on page 165
- Configuration Examples for MVR, on page 167

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.

Guidelines and Limitations for MVR

MVR has the following guidelines and limitations:

- MVR is supported only for Cisco Nexus 9508 switches with N9K-X9636C-R, N9K-X9636C-RX, or N9K-X9636Q-R line cards.
- 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 21: 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.

SUMMARY STEPS

1. **configure terminal**
2. [**no**]**mvr**
3. [**no**] **mvr-vlan** *vlan-id*
4. [**no**] **mvr-group** *addr* [/*mask*] [**count** *groups*] [**vlan** *vlan-id*]
5. (Optional) **clear mvr counters** [**source-ports** | **receiver-ports**]
6. (Optional) **show mvr**
7. (Optional) **copy running-config startup-config**

DETAILED STEPS

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 Cisco NX-OS device.

SUMMARY STEPS

1. **configure terminal**
2. **mvr**
3. **interface** {**ethernet** *slot/port* | **port-channel** *channel-number* | **vethernet** *number*}
4. [**no**] **mvr-type** {**source** | **receiver**}
5. (Optional) [**no**] **mvr-vlan** *vlan-id*
6. (Optional) [**no**] **mvr-group** *addr* [*/mask*] [**vlan** *vlan-id*]
7. (Optional) **copy running-config startup-config**

DETAILED STEPS

Procedure

	Command or Action	Purpose

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

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:**	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
	`switch(config-mvr-if)# mvr-group 225.1.3.1 vlan 100`	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.

SUMMARY STEPS

1. **configure terminal**
2. **mvr-config**
3. **mvr-suppress-query vlan** *vlan-ID*

DETAILED 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

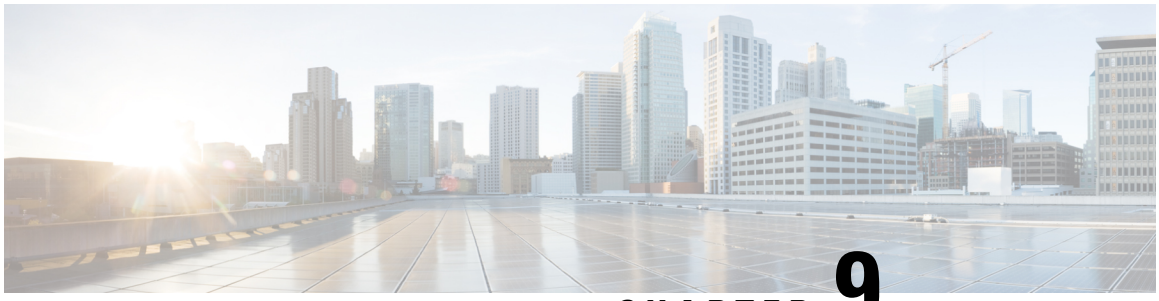
The following example shows how to globally enable MVR and configure the global parameters:

```
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 9

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), on page 169
- Guidelines and Limitations for NLB, on page 170
- Prerequisites for Microsoft Network Load Balancing (NLB), on page 171
- Multicast Mode, on page 171
- IGMP Multicast Mode, on page 172
- Verifying the NLB Configuration, on page 173

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 171
- **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 172 for configuration examples.

This section describes how to configure a Cisco Nexus 9000 series 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.

Guidelines and Limitations for NLB

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

- Beginning with Cisco NX-OS Release 9.3(5), Multicast NLB is supported on Cisco Nexus 9300-FX3 platform switches.
- Multicast NLB is supported on Cisco Nexus 9300-EX, Cisco Nexus 9300-FX, Nexus 9300-FX2 platform switches, Cisco Nexus 9500 platform switches with N9K-X9700-EX line cards, N9K-X9700-FX line cards, Cisco Nexus 9500 platform switches with N9K-C9500-FM-E fabric cards and N9K-C9500-FM-E2 fabric cards. Beginning with Cisco NX-OS Release 9.3(6), Multicast NLB is supported on Cisco Nexus 9300-GX platform switches.
  - Multicast NLB is not supported on the Cisco Nexus 9500 modules with N9K-C9508-FM-2.
  - Multicast NLB is not supported on the Cisco Nexus 9300 and 9364C switches.
  - L2 (switched multicast) and L3 (routed multicast) is not supported to, from or inside of a VLAN that is configured for multicast NLB. This includes link local multicast groups as well, thus control plane protocols that use these groups are not supported to be configured on these VLANs.
  - Note that HSRP and VRRP are not included in the above mentioned limitations.
- Flooding for Microsoft Network Load Balancing (NLB) unicast mode is not supported on Cisco Nexus 9000 switches. A static ARP entry must be configured to map the NLB virtual IP address to the NLB virtual MAC address. Furthermore, a static MAC address entry must be configured to map the NLB virtual MAC address to a specific egress interface.
- 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.
- NLB in the unicast, multicast, and IGMP multicast modes is not supported on Cisco Nexus 9000 Series based VXLAN VTEPs. The work around is to move the NLB cluster behind intermediary device (which supports NLB in the respective mode) and inject the cluster IP address as external prefix into VXLAN fabric.

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.

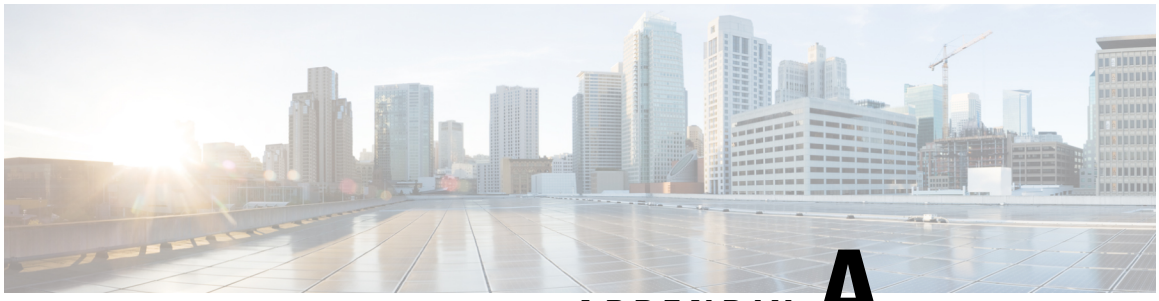
3. Configure static IGMP snooping group entries for the interfaces connected to the NLB server that needs the traffic:

```
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.

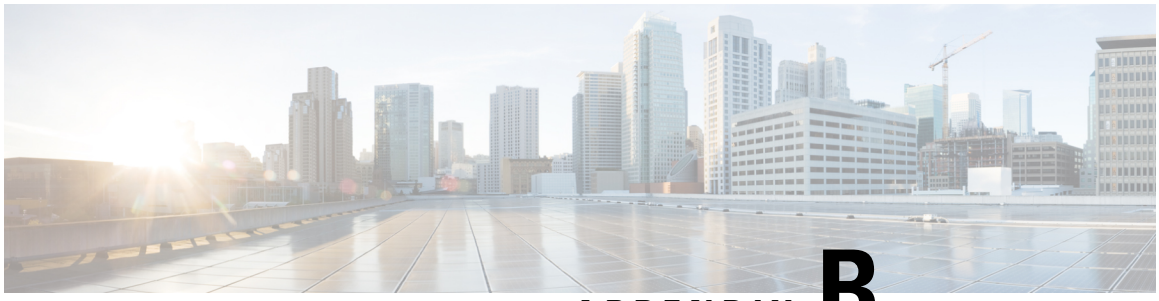
- IETF RFCs for IP Multicast, on page 175

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 2858	*Multiprotocol Extensions for BGP-4*
RFC 3376	*Internet Group Management Protocol*
RFC 3446	*Anycast Rendezvous Point (RP) mechanism using Protoco Multicast (PIM) and Multicast Source Discovery Protoco*
RFC 3618	*Multicast Source Discovery Protocol (MSDP)*
RFC 4601	*Protocol Independent Multicast - Sparse Mode (PIM-SM) 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 Cisco NX-OS Multicast

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

- Configuration Limits, on page 177

Configuration Limits

The features supported by Cisco 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 Cisco Nexus 9000 Series NX-OS Verified Scalability Guide.

Configuration Limits for Cisco NX-OS Multicast