Multicast Configuration Guide for the Cisco NCS 6000 Series Routers, IOS XR Release 7.1.x

First Published: 2020-01-29

Last Modified: 2023-06-05

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

CONTENTS

PREFACE

CHAPTER 1

CHAPTER 2

Preface

Note This product has reached end-of-life status. For more information, see the End-of-Life and End-of-Sale Notices.

The preface contains these sections:

Changes to This Document

Describes the changes in the document from the initial release of this document.

Table 1: Changes to This Document

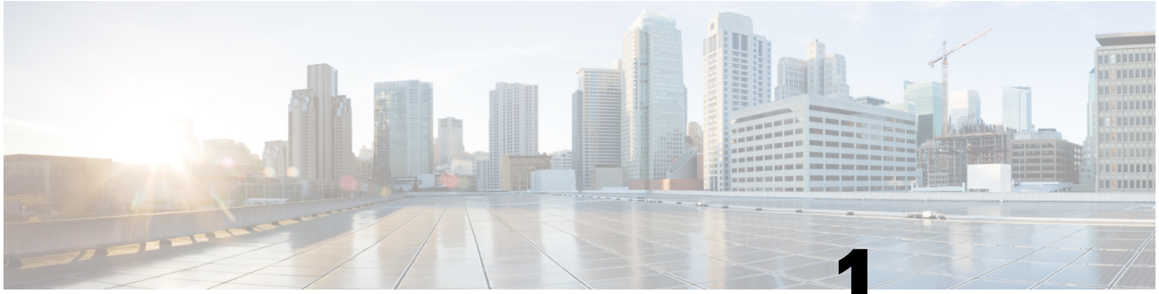
Date	**Summary**
January 2020	Initial release of this document.

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 Marketplace.
- 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 web-based tool that acts as a gateway to the Cisco bug tracking system that maintains a comprehensive list of defects and vulnerabilities in Cisco products and software. BST provides you with detailed defect information about your products and software.



CHAPTER 1

New and Changed Multicast Features

This chapter lists all the features that have been added or modified in this guide. The table also contains references to these feature documentation sections.

Multicast Features Added or Modified in IOS XR Release 7.1.x

Table 2: New and Changed Features

Feature	Description	Changed in Release	Where Documented
MVPN GRE over PWHE with CSI	MVPN GRE over PWHE is supported on CSI interface	Release 7.0.1	MVPN GRE over PWHE with CSI, on page 54

CHAPTER 2

Implementing Multicast Routing on Cisco IOS XR Software

Multicast routing is a bandwidth-conserving technology that reduces traffic by simultaneously delivering a single stream of information to potentially thousands of corporate recipients and homes. Applications that take advantage of multicast routing include video conferencing, corporate communications, distance learning, and distribution of software, stock quotes, and news.

This document assumes that you are familiar with IPv4 multicast routing configuration tasks and concepts for Cisco IOS XR Software .

Multicast routing allows a host to send packets to a subset of all hosts as a group transmission rather than to a single host, as in unicast transmission, or to all hosts, as in broadcast transmission. The subset of hosts is known as **group members** and are identified by a single multicast group address that falls under the IP Class D address range from 224.0.0.0 through 239.255.255.255.

Note NCS 6000 does not support multicast over PWHE interfaces.

Feature History for Configuring Multicast Routing on the Cisco NCS 6000 Series Routers

Release	Modification
Release 5.0.0	This feature was introduced.
Release 6.1.2	Point-to-Multipoint Traffic Engineering with GTM feature was introduced.
Release 6.1.2	MLDP Edge feature was introduced.

Prerequisites for Implementing Multicast Routing

- You must install and activate the multicast package.
- You must be in a user group associated with a task group that includes the proper task IDs. The command reference guides include the task IDs required for each command. If you suspect user group assignment is preventing you from using a command, contact your AAA administrator for assistance.
- You must be familiar with IPv4 multicast routing configuration tasks and concepts.
- Unicast routing must be operational.

Note PW-EH interface is supported from Cisco IOS XR, Release 6.3.1. Multicast Routing is not supported on the PW-EH interface for now.

If below configuration in Example 1 is enabled, please disable multicast routing in PW-EH interface manually using the CLI commands in Example 2.

Example 1:

```
multicast-routing
address-family ipv4
 interface all enable
!
```

Example 2:

```
  interface PW-Ether1
   disable
```

Information About Implementing Multicast Routing

Key Protocols and Features Supported in the Cisco IOS XR Software Multicast Routing Implementation

Multicast Routing Functional Overview

Traditional IP communication allows a host to send packets to a single host (*unicast transmission*) or to all hosts (*broadcast transmission*). Multicast provides a third scheme, allowing a host to send a single data stream to a subset of all hosts (*group transmission*) at about the same time. IP hosts are known as group members.

Packets delivered to group members are identified by a single multicast group address. Multicast packets are delivered to a group using best-effort reliability, just like IP unicast packets.

The multicast environment consists of senders and receivers. Any host, regardless of whether it is a member of a group, can send to a group. However, only the members of a group receive the message.

A multicast address is chosen for the receivers in a multicast group. Senders use that group address as the destination address of a datagram to reach all members of the group.

Membership in a multicast group is dynamic; hosts can join and leave at any time. There is no restriction on the location or number of members in a multicast group. A host can be a member of more than one multicast group at a time.

How active a multicast group is and what members it has can vary from group to group and from time to time. A multicast group can be active for a long time, or it may be very short-lived. Membership in a group can change constantly. A group that has members may have no activity.

Many multimedia applications involve multiple participants. Multicast is naturally suitable for this communication paradigm.

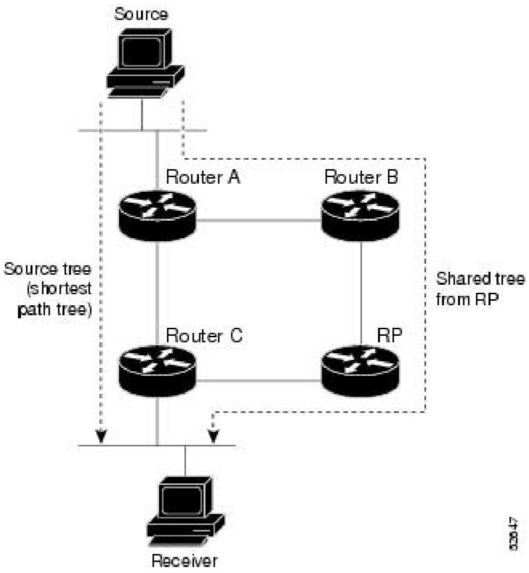
Multicast Routing Implementation

Cisco IOS XR Software supports the following protocols to implement multicast routing:

- IGMP used between hosts on a LAN and the routers on that LAN to track the multicast groups of which hosts are members.
- Protocol Independent Multicast in sparse mode (PIM-SM) is used between routers so that they can track which multicast packets to forward to each other and to their directly connected LANs.
- Protocol Independent Multicast in Source-Specific Multicast (PIM-SSM) is similar to PIM-SM with the additional ability to report interest in receiving packets from specific source addresses (or from all but the specific source addresses), to an IP multicast address.
- PIM Bidirectional is a variant of the Protocol Independent Multicast suit of routing protocols for IP multicast. PIM-BIDIR is designed to be used for many-to-many applications within individual PIM domains.

This image shows IGMP and PIM-SM operating in a multicast environment.

Figure 1: Multicast Routing Protocols



PIM-SM, PIM-SSM, and PIM-BIDIR

Protocl Independent Multicast (PIM) is a multicast routing protocol used to create multicast distribution trees, which are used to forward multicast data packets. PIM is an efficient IP routing protocol that is “independent” of a routing table, unlike other multicast protocols such as Multicast Open Shortest Path First (MOSPF) or Distance Vector Multicast Routing Protocol (DVMRP).

Cisco IOS XR Software supports Protocol Independent Multicast in sparse mode (PIM-SM), Protocol Independent Multicast in Source-Specific Multicast (PIM-SSM), and Protocol Independent Multicast in Bi-directional mode (BIDIR) permitting these modes to operate on your router at the same time.

PIM-SM and PIM-SSM supports one-to-many applications by greatly simplifying the protocol mechanics for deployment ease. Bidir PIM helps deploy emerging communication and financial applications that rely on a many-to-many applications model. BIDIR PIM enables these applications by allowing them to easily scale to a very large number of groups and sources by eliminating the maintenance of source state.

PIM-SM Operations

PIM in sparse mode operation is used in a multicast network when relatively few routers are involved in each multicast and these routers do not forward multicast packets for a group, unless there is an explicit request for the traffic.

For more information about PIM-SM, see the PIM-Sparse Mode, on page 9.

PIM-SSM Operations

PIM in Source-Specific Multicast operation uses information found on source addresses for a multicast group provided by receivers and performs source filtering on traffic.

- By default, PIM-SSM operates in the 232.0.0.0/8 multicast group range for IPv4 . To configure these values, use the **ssm range** command.

- If SSM is deployed in a network already configured for PIM-SM, only the last-hop routers must be upgraded with Cisco IOS XR Software that supports the SSM feature.
- No MSDP SA messages within the SSM range are accepted, generated, or forwarded.

Restrictions for PIM-SM and PIM-SSM, and PIM BIDIR

Interoperability with SSM

PIM-SM operations within the SSM range of addresses change to PIM-SSM. In this mode, only PIM (S,G) join and prune messages are generated by the router, and no (S,G) RP shared tree or (*,G) shared tree messages are generated.

IGMP Version

To report multicast memberships to neighboring multicast routers, hosts use IGMP, and all routers on the subnet must be configured with the same version of IGMP.

A router running Cisco IOS XR Software does not automatically detect Version 1 systems. You must use the **version** command in router IGMP configuration submode to configure the IGMP version.

Internet Group Management Protocol

Cisco IOS XR Software provides support for Internet Group Management Protocol (IGMP) over IPv4

IGMP a means for hosts to indicate which multicast traffic they are interested in and for routers to control and limit the flow of multicast traffic throughout the network. Routers build state by means of IGMP messages; that is, router queries and host reports.

A set of queries and hosts that receive multicast data streams from the same source is called a *multicast group*. Hosts use IGMP messages to join and leave multicast groups.

Note IGMP messages use group addresses, which are Class D IP addresses. The high-order four bits of a Class D address are 1110. Host group addresses can be in the range 224.0.0.0 to 239.255.255.255. The address 224.0.0.0 is guaranteed not to be assigned to any group. The address 224.0.0.1 is assigned to all systems on a subnet. The address 224.0.0.2 is assigned to all routers on a subnet.

IGMP Versions

The following points describe IGMP versions 1, 2, and 3:

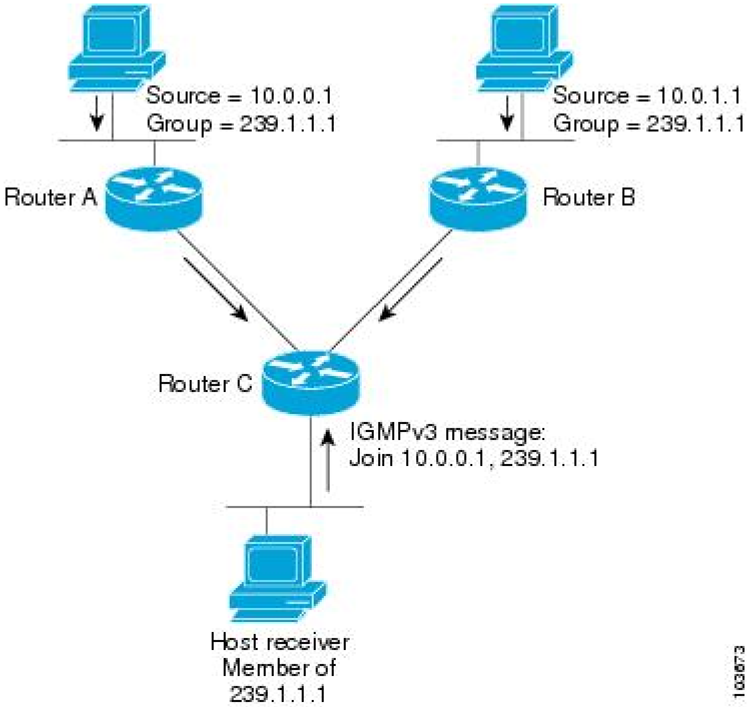
- IGMP Version 1 provides for the basic query-response mechanism that allows the multicast router to determine which multicast groups are active and for other processes that enable hosts to join and leave a multicast group.
- IGMP Version 2 extends IGMP allowing such features as the IGMP query timeout and the maximum query-response time. See RFC 2236.
- IGMP Version 3 permits joins and leaves for certain source and group pairs instead of requesting traffic from all sources in the multicast group.

IGMP Routing Example

Figure 2: IGMPv3 Signaling, on page 8 illustrates two sources, 10.0.0.1 and 10.0.1.1, that are multicasting to group 239.1.1.1. The receiver wants to receive traffic addressed to group 239.1.1.1 from source 10.0.0.1 but not from source 10.0.1.1. The host must send an IGMPv3 message containing a list of sources and groups (S, G) that it wants to join and a list of sources and groups (S, G) that it wants to leave. Router C can now use this information to prune traffic from Source 10.0.1.1 so that only Source 10.0.0.1 traffic is being delivered to

Router C.

Figure 2: IGMPv3 Signaling



Note When configuring IGMP, ensure that all systems on the subnet support the same IGMP version. The router does not automatically detect Version 1 systems. Configure the router for Version 2 if your hosts do not support Version 3.

Protocol Independent Multicast

Protocol Independent Multicast (PIM) is a routing protocol designed to send and receive multicast routing updates. Proper operation of multicast depends on knowing the unicast paths towards a source or an RP. PIM relies on unicast routing protocols to derive this reverse-path forwarding (RPF) information. As the name PIM implies, it functions independently of the unicast protocols being used. PIM relies on the Routing Information Base (RIB) for RPF information.

The Cisco IOS XR implementation of PIM is based on RFC 4601 Protocol Independent Multicast - Sparse Mode (PIM-SM): Protocol Specification. For more information, see RFC 4601 and the Protocol Independent Multicast (PIM): Motivation and Architecture Internet Engineering Task Force (IETF) Internet draft.

Note Cisco IOS XR Software supports PIM-SM, PIM-SSM, and PIM Version 2 only. PIM Version 1 hello messages that arrive from neighbors are rejected.

PIM-Sparse Mode

Typically, PIM in sparse mode (PIM-SM) operation is used in a multicast network when relatively few routers are involved in each multicast. Routers do not forward multicast packets for a group, unless there is an explicit request for traffic. Requests are accomplished using PIM join messages, which are sent hop by hop toward the root node of the tree. The root node of a tree in PIM-SM is the rendezvous point (RP) in the case of a shared tree or the first-hop router that is directly connected to the multicast source in the case of a shortest path tree (SPT). The RP keeps track of multicast groups, and the sources that send multicast packets are registered with the RP by the first-hop router of the source.

As a PIM join travels up the tree, routers along the path set up the multicast forwarding state so that the requested multicast traffic is forwarded back down the tree. When multicast traffic is no longer needed, a router sends a PIM prune message up the tree toward the root node to prune (or remove) the unnecessary traffic. As this PIM prune travels hop by hop up the tree, each router updates its forwarding state appropriately. Ultimately, the forwarding state associated with a multicast group or source is removed. Additionally, if prunes are not explicitly sent, the PIM state will timeout and be removed in the absence of any further join messages.

PIM-SM is the best choice for multicast networks that have potential members at the end of WAN links.

PIM-Source Specific Multicast

In many multicast deployments where the source is known, protocol-independent multicast-source-specific multicast (PIM-SSM) mapping is the obvious multicast routing protocol choice to use because of its simplicity. Typical multicast deployments that benefit from PIM-SSM consist of entertainment-type solutions like the ETTH space, or financial deployments that completely rely on static forwarding.

PIM-SSM is derived from PIM-SM. However, whereas PIM-SM allows for the data transmission of all sources sending to a particular group in response to PIM join messages, the SSM feature forwards traffic to receivers only from those sources that the receivers have explicitly joined. Because PIM joins and prunes are sent directly towards the source sending traffic, an RP and shared trees are unnecessary and are disallowed. SSM is used to optimize bandwidth utilization and deny unwanted Internet broadcast traffic. The source is provided by interested receivers through IGMPv3 membership reports.

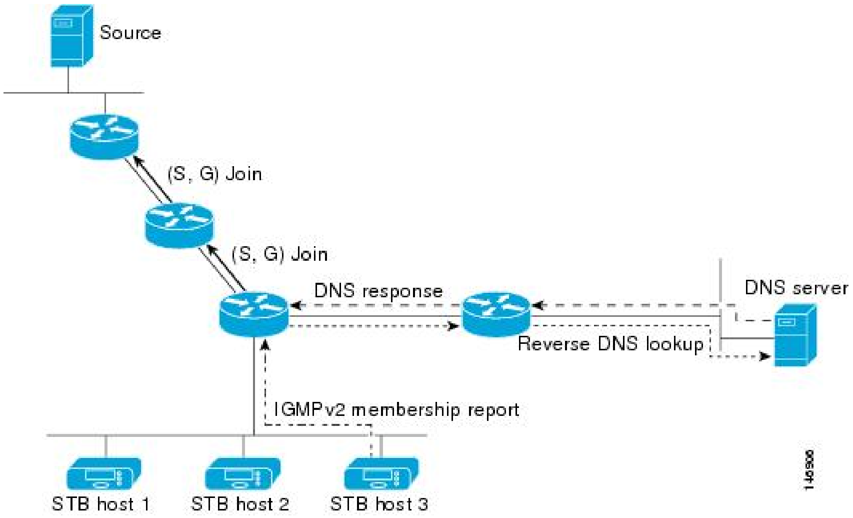
In SSM, delivery of datagrams is based on (S,G) channels. Traffic for one (S,G) channel consists of datagrams with an IP unicast source address S and the multicast group address G as the IP destination address. Systems receive traffic by becoming members of the (S,G) channel. Signaling is not required, but receivers must subscribe or unsubscribe to (S,G) channels to receive or not receive traffic from specific sources. Channel subscription signaling uses IGMP to include mode membership reports, which are supported only in Version 3 of IGMP (IGMPv3).

To run SSM with IGMPv3, SSM must be supported on the multicast router, the host where the application is running, and the application itself. Cisco IOS XR Software allows SSM configuration for an arbitrary subset of the IP multicast address range 224.0.0.0 through 239.255.255.255. When an SSM range is defined, existing IP multicast receiver applications do not receive any traffic when they try to use addresses in the SSM range, unless the application is modified to use explicit (S,G) channel subscription.

DNS-based SSM Mapping

DNS-based SSM mapping enables you to configure the last hop router to perform a reverse DNS lookup to determine sources sending to groups (see the figure below). When DNS-based SSM mapping is configured, the router constructs a domain name that includes the group address G and performs a reverse lookup into the DNS. The router looks up IP address resource records (IP A RRs) to be returned for this constructed domain name and uses the returned IP addresses as the source addresses associated with this group. SSM mapping supports up to 20 sources for each group. The router joins all sources configured for a group.

Figure 3: DNS-based SSM Mapping



The SSM mapping mechanism that enables the last hop router to join multiple sources for a group can be used to provide source redundancy for a TV broadcast. In this context, the redundancy is provided by the last hop router using SSM mapping to join two video sources simultaneously for the same TV channel. However, to prevent the last hop router from duplicating the video traffic, it is necessary that the video sources utilize a server-side switchover mechanism where one video source is active while the other backup video source is passive. The passive source waits until an active source failure is detected before sending the video traffic for the TV channel. The server-side switchover mechanism, thus, ensures that only one of the servers is actively sending the video traffic for the TV channel.

To look up one or more source addresses for a group G that includes G1, G2, G3, and G4, the following DNS resource records (RRs) must be configured on the DNS server:

G4.G3.G2.G1 [*multicast-domain*] [*timeout*]	IN A *source-address-1*
	IN A *source-address-2*
	IN A *source-address-n*

The *multicast-domain* argument is a configurable DNS prefix. The default DNS prefix is in-addr.arpa. You should only use the default prefix when your installation is either separate from the internet or if the group names that you map are global scope group addresses (RFC 2770 type addresses that you configure for SSM) that you own.

The *timeout* argument configures the length of time for which the router performing SSM mapping will cache the DNS lookup. This argument is optional and defaults to the timeout of the zone in which this entry is configured. The timeout indicates how long the router will keep the current mapping before querying the DNS server for this group. The timeout is derived from the cache time of the DNS RR entry and can be configured for each group/source entry on the DNS server. You can configure this time for larger values if you want to minimize the number of DNS queries generated by the router. Configure this time for a low value if you want to be able to quickly update all routers with new source addresses.

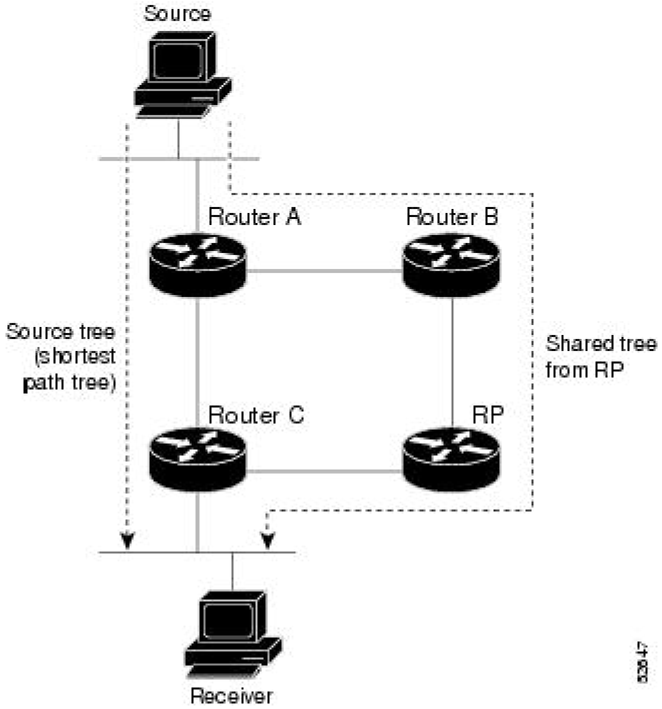
Note See your DNS server documentation for more information about configuring DNS RRs.

To configure DNS-based SSM mapping in the software, you must configure a few global commands but no per-channel specific configuration is needed. There is no change to the configuration for SSM mapping if additional channels are added. When DNS-based SSM mapping is configured, the mappings are handled entirely by one or more DNS servers. All DNS techniques for configuration and redundancy management can be applied to the entries needed for DNS-based SSM mapping.

PIM Shared Tree and Source Tree (Shortest Path Tree)

In PIM-SM, the rendezvous point (RP) is used to bridge sources sending data to a particular group with receivers sending joins for that group. In the initial setup of state, interested receivers receive data from senders to the group across a single data distribution tree rooted at the RP. This type of distribution tree is called a shared tree or rendezvous point tree (RPT) as illustrated in Figure 4: Shared Tree and Source Tree (Shortest Path Tree), on page 11 . Data from senders is delivered to the RP for distribution to group members joined to the shared tree.

Figure 4: Shared Tree and Source Tree (Shortest Path Tree)



Unless the **spt-threshold infinity** command is configured, this initial state gives way as soon as traffic is received on the leaf routers (designated router closest to the host receivers). When the leaf router receives traffic from the RP on the RPT, the router initiates a switch to a data distribution tree rooted at the source sending traffic. This type of distribution tree is called a **shortest path tree** or **source tree**. By default, the Cisco IOS XR Software switches to a source tree when it receives the first data packet from a source.

The following process describes the move from shared tree to source tree in more detail:

1. Receiver joins a group; leaf Router C sends a join message toward RP.
2. RP puts link to Router C in its outgoing interface list.
3. Source sends data; Router A encapsulates data in Register and sends it to RP.
4. RP forwards data down the shared tree to Router C and sends a join message toward Source. At this point, data may arrive twice at the RP, once encapsulated and once natively.
5. When data arrives natively (unencapsulated) at RP, RP sends a register-stop message to Router A.
6. By default, receipt of the first data packet prompts Router C to send a join message toward Source.
7. When Router C receives data on (S,G), it sends a prune message for Source up the shared tree.
8. RP deletes the link to Router C from outgoing interface of (S,G). RP triggers a prune message toward Source.

Join and prune messages are sent for sources and RPs. They are sent hop by hop and are processed by each PIM router along the path to the source or RP. Register and register-stop messages are not sent hop by hop. They are exchanged using direct unicast communication between the designated router that is directly connected to a source and the RP for the group.

Tip The **spt-threshold infinity** command lets you configure the router so that it never switches to the shortest path tree (SPT).

Multicast-Intact

The multicast-intact feature provides the ability to run multicast routing (PIM) when Interior Gateway Protocol (IGP) shortcuts are configured and active on the router. Both Open Shortest Path First, version 2 (OSPFv2), and Intermediate System-to-Intermediate System (IS-IS) support the multicast-intact feature. Multiprotocol Label Switching Traffic Engineering (MPLS-TE) and IP multicast coexistence is supported in Cisco IOS XR Software by using the **mpls traffic-eng multicast-intact** IS-IS or OSPF router command. See *Routing Configuration Guide for Cisco NCS 6000 Series Routers* for information on configuring multicast intact using IS-IS and OSPF commands.

You can enable multicast-intact in the IGP when multicast routing protocols (PIM) are configured and IGP shortcuts are configured on the router. IGP shortcuts are MPLS tunnels that are exposed to IGP. The IGPs route the IP traffic over these tunnels to destinations that are downstream from the egress router of the tunnel (from an SPF perspective). PIM cannot use IGP shortcuts for propagating PIM joins because reverse path forwarding (RPF) cannot work across a unidirectional tunnel.

When you enable multicast-intact on an IGP, the IGP publishes a parallel or alternate set of equal-cost next-hops for use by PIM. These next-hops are called **mcast-intact next-hops**. The mcast-intact next-hops have the following attributes:

- They are guaranteed not to contain any IGP shortcuts.

- They are not used for unicast routing but are used only by PIM to look up an IPv4 next hop to a PIM source.
- They are not published to the Forwarding Information Base (FIB).
- When multicast-intact is enabled on an IGP, all IPv4 destinations that were learned through link-state advertisements are published with a set equal-cost mcast-intact next-hops to the RIB. This attribute applies even when the native next-hops have no IGP shortcuts.
- In IS-IS, the max-paths limit is applied by counting both the native and mcast-intact next-hops together. (In OSPFv2, the behavior is slightly different.)

Designated Routers

Cisco routers use PIM-SM to forward multicast traffic and follow an election process to select a designated router (DR) when there is more than one router on a LAN segment.

The designated router is responsible for sending PIM register and PIM join and prune messages toward the RP to inform it about host group membership.

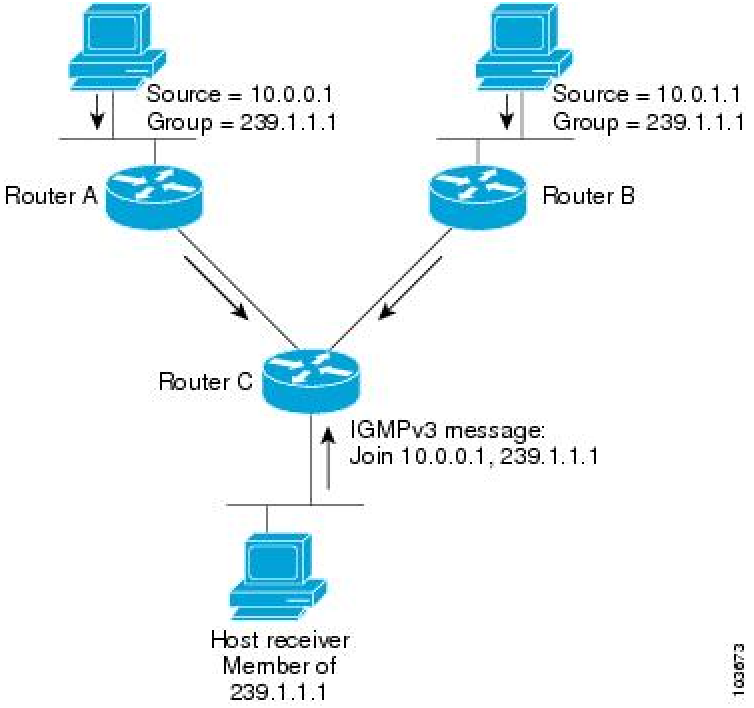
If there are multiple PIM-SM routers on a LAN, a designated router must be elected to avoid duplicating multicast traffic for connected hosts. The PIM router with the highest IP address becomes the DR for the LAN unless you choose to force the DR election by use of the **dr-priority** command. The DR priority option allows you to specify the DR priority of each router on the LAN segment (default priority = 1) so that the router with the highest priority is elected as the DR. If all routers on the LAN segment have the same priority, the highest IP address is again used as the tiebreaker.

Figure 5: Designated Router Election on a Multiaccess Segment, on page 14 illustrates what happens on a multiaccess segment. Router A (10.0.0.253) and Router B (10.0.0.251) are connected to a common multiaccess Ethernet segment with Host A (10.0.0.1) as an active receiver for Group A. As the Explicit Join model is used, only Router A, operating as the DR, sends joins to the RP to construct the shared tree for Group A. If Router B were also permitted to send (*, G) joins to the RP, parallel paths would be created and Host A would receive duplicate multicast traffic. When Host A begins to source multicast traffic to the group, the DR's responsibility is to send register messages to the RP. Again, if both routers were assigned the responsibility, the RP would receive duplicate multicast packets.

If the DR fails, the PIM-SM provides a way to detect the failure of Router A and to elect a failover DR. If the DR (Router A) were to become inoperable, Router B would detect this situation when its neighbor adjacency with Router A timed out. Because Router B has been hearing IGMP membership reports from Host A, it already has IGMP state for Group A on this interface and immediately sends a join to the RP when it becomes the new DR. This step reestablishes traffic flow down a new branch of the shared tree using Router B. Additionally, if Host A were sourcing traffic, Router B would initiate a new register process immediately after receiving the next multicast packet from Host A. This action would trigger the RP to join the SPT to Host A, using a new branch through Router B.

Tip Two PIM routers are neighbors if there is a direct connection between them. To display your PIM neighbors, use the **show pim neighbor** command in EXEC mode.

Figure 5: Designated Router Election on a Multiaccess Segment



Note DR election process is required only on multiaccess LANs. The last-hop router directly connected to the host is the DR.

Rendezvous Points

When PIM is configured in sparse mode, you must choose one or more routers to operate as a rendezvous point (RP). A rendezvous point is a single common root placed at a chosen point of a shared distribution tree, as illustrated in Figure 4: Shared Tree and Source Tree (Shortest Path Tree), on page 11. A rendezvous point can be either configured statically in each box or learned through a dynamic mechanism.

PIM DRs forward data from directly connected multicast sources to the rendezvous point for distribution down the shared tree. Data is forwarded to the rendezvous point in one of two ways:

- Encapsulated in register packets and unicast directly to the rendezvous point by the first-hop router operating as the DR
- Multicast forwarded by the RPF forwarding algorithm, described in the Reverse-Path Forwarding, on page 16, if the rendezvous point has itself joined the source tree.

The rendezvous point address is used by first-hop routers to send PIM register messages on behalf of a host sending a packet to the group. The rendezvous point address is also used by last-hop routers to send PIM join and prune messages to the rendezvous point to inform it about group membership. You must configure the rendezvous point address on all routers (including the rendezvous point router).

A PIM router can be a rendezvous point for more than one group. Only one rendezvous point address can be used at a time within a PIM domain. The conditions specified by the access list determine for which groups the router is a rendezvous point.

You can either manually configure a PIM router to function as a rendezvous point or allow the rendezvous point to learn group-to-RP mappings automatically by configuring Auto-RP or BSR. (For more information, see the Auto-RP, on page 15 section that follows and PIM Bootstrap Router, on page 15.)

Auto-RP

Automatic route processing (Auto-RP) is a feature that automates the distribution of group-to-RP mappings in a PIM network. This feature has these benefits:

- It is easy to use multiple RPs within a network to serve different group ranges.
- It allows load splitting among different RPs.
- It facilitates the arrangement of RPs according to the location of group participants.
- It avoids inconsistent, manual RP configurations that might cause connectivity problems.

Multiple RPs can be used to serve different group ranges or to serve as hot backups for each other. To ensure that Auto-RP functions, configure routers as candidate RPs so that they can announce their interest in operating as an RP for certain group ranges. Additionally, a router must be designated as an RP-mapping agent that receives the RP-announcement messages from the candidate RPs, and arbitrates conflicts. The RP-mapping agent sends the consistent group-to-RP mappings to all remaining routers. Thus, all routers automatically determine which RP to use for the groups they support.

Tip By default, if a given group address is covered by group-to-RP mappings from both static RP configuration, and is discovered using Auto-RP or PIM BSR, the Auto-RP or PIM BSR range is preferred. To override the default, and use only the RP mapping, use the **rp-address override** keyword.

PIM Bootstrap Router

The PIM bootstrap router (BSR) provides a fault-tolerant, automated RP discovery and distribution mechanism that simplifies the Auto-RP process. This feature is enabled by default allowing routers to dynamically learn the group-to-RP mappings.

PIM uses the BSR to discover and announce RP-set information for each group prefix to all the routers in a PIM domain. This is the same function accomplished by Auto-RP, but the BSR is part of the PIM Version 2 specification. The BSR mechanism interoperates with Auto-RP on Cisco routers.

To avoid a single point of failure, you can configure several candidate BSRs in a PIM domain. A BSR is elected among the candidate BSRs automatically. Candidates use bootstrap messages to discover which BSR has the highest priority. The candidate with the highest priority sends an announcement to all PIM routers in the PIM domain that it is the BSR.

Routers that are configured as candidate RPs unicast to the BSR the group range for which they are responsible. The BSR includes this information in its bootstrap messages and disseminates it to all PIM routers in the domain. Based on this information, all routers are able to map multicast groups to specific RPs. As long as a router is receiving the bootstrap message, it has a current RP map.

Reverse-Path Forwarding

Reverse-path forwarding (RPF) is an algorithm used for forwarding multicast datagrams. It functions as follows:

- If a router receives a datagram on an interface it uses to send unicast packets to the source, the packet has arrived on the RPF interface.
- If the packet arrives on the RPF interface, a router forwards the packet out the interfaces present in the outgoing interface list of a multicast routing table entry.
- If the packet does not arrive on the RPF interface, the packet is silently discarded to prevent loops.

PIM uses both source trees and RP-rooted shared trees to forward datagrams; the RPF check is performed differently for each, as follows:

- If a PIM router has an (S,G) entry present in the multicast routing table (a source-tree state), the router performs the RPF check against the IP address of the source for the multicast packet.
- If a PIM router has no explicit source-tree state, this is considered a shared-tree state. The router performs the RPF check on the address of the RP, which is known when members join the group.

Sparse-mode PIM uses the RPF lookup function to determine where it needs to send joins and prunes. (S,G) joins (which are source-tree states) are sent toward the source. (*,G) joins (which are shared-tree states) are sent toward the RP.

Multicast VPN

Multicast VPN (MVPN) provides the ability to dynamically provide multicast support over MPLS networks. MVPN introduces an additional set of protocols and procedures that help enable a provider to support multicast traffic in a VPN.

Note PIM-Bidir is not supported on MVPN.

There are two ways MCAST VPN traffic can be transported over the core network:

- Rosen GRE (native): MVPN uses GRE with unique multicast distribution tree (MDT) forwarding to enable scalability of native IP Multicast in the core network. MVPN introduces multicast routing information to the VPN routing and forwarding table (VRF), creating a Multicast VRF. In Rosen GRE, the MCAST customer packets (c-packets) are encapsulated into the provider MCAST packets (p-packets), so that the PIM protocol is enabled in the provider core, and mrib/mfib is used for forwarding p-packets in the core.
- MLDP ones (Rosen, partition): MVPN allows a service provider to configure and support multicast traffic in an MPLS VPN environment. This type supports routing and forwarding of multicast packets for each individual VPN routing and forwarding (VRF) instance, and it also provides a mechanism to transport VPN multicast packets across the service provider backbone. In the MLDP case, the regular label switch path forwarding is used, so core does not need to run PIM protocol. In this scenario, the c-packets are encapsulated in the MPLS labels and forwarding is based on the MPLS Label Switched Paths (LSPs) ,similar to the unicast case.

In both the above types, the MVPN service allows you to build a Protocol Independent Multicast (PIM) domain that has sources and receivers located in different sites.

To provide Layer 3 multicast services to customers with multiple distributed sites, service providers look for a secure and scalable mechanism to transmit customer multicast traffic across the provider network. Multicast VPN (MVPN) provides such services over a shared service provider backbone, using native multicast technology similar to BGP/MPLS VPN.

MVPN emulates MPLS VPN technology in its adoption of the multicast domain (MD) concept, in which provider edge (PE) routers establish virtual PIM neighbor connections with other PE routers that are connected to the same customer VPN. These PE routers thereby form a secure, virtual multicast domain over the provider network. Multicast traffic is then transmitted across the core network from one site to another, as if the traffic were going through a dedicated provider network.

Multi-instance BGP is supported on multicast and MVPN. Multicast-related SAFIs can be configured on multiple BGP instances.

Multicast VPN Routing and Forwarding

Dedicated multicast routing and forwarding tables are created for each VPN to separate traffic in one VPN from traffic in another.

The VPN-specific multicast routing and forwarding database is referred to as **MVRF**. On a PE router, an MVRF is created when multicast is enabled for a VRF. Protocol Independent Multicast (PIM), and Internet Group Management Protocol (IGMP) protocols run in the context of MVRF, and all routes created by an MVRF protocol instance are associated with the corresponding MVRF. In addition to VRFs, which hold VPN-specific protocol states, a PE router always has a global VRF instance, containing all routing and forwarding information for the provider network.

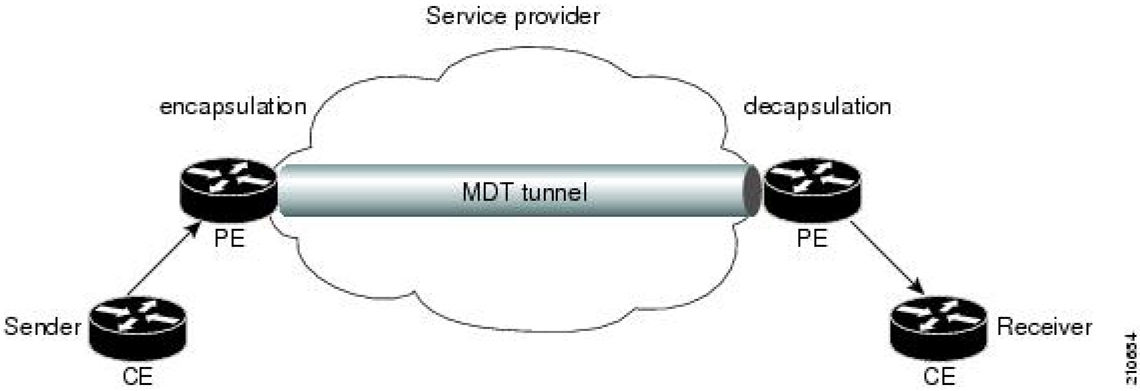
Multicast Distribution Tree Tunnels

The multicast distribution tree (MDT) can span multiple customer sites through provider networks, allowing traffic to flow from one source to multiple receivers. For MLDP, the MDT tunnel are called Labeled MDT (LMDT).

Secure data transmission of multicast packets sent from the customer edge (CE) router at the ingress PE router is achieved by encapsulating the packets in a provider header and transmitting the packets across the core. At the egress PE router, the encapsulated packets are decapsulated and then sent to the CE receiving routers.

Multicast distribution tree (MDT) tunnels are point-to-multipoint. A MDT tunnel interface is an interface that MVRF uses to access the multicast domain. It can be deemed as a passage that connects an MVRF and the global MVRF. Packets sent to an MDT tunnel interface are received by multiple receiving routers. Packets sent to an MDT tunnel interface are encapsulated, and packets received from a MDT tunnel interface are decapsulated.

Figure 6: Virtual PIM Peer Connection over an MDT Tunnel Interface



Encapsulating multicast packets in a provider header allows PE routers to be kept unaware of the packets' origin—all VPN packets passing through the provider network are viewed as native multicast packets and are routed based on the routing information in the core network. To support MVPN, PE routers only need to support native multicast routing.

MVPN also supports optimized VPN traffic forwarding for high-bandwidth applications that have sparsely distributed receivers. A dedicated multicast group can be used to encapsulate packets from a specific source, and an optimized MDT can be created to send traffic only to PE routers connected to interested receivers. This is referred to **data MDT**.

The rate at which the MVPN traffic switches over from the default to data MDT is 1/*n* th the configured threshold value, where *n* represents the number of slices in an incoming line card.

InterAS Support on Multicast VPN

The Multicast VPN Inter-AS Support feature enables service providers to provide multicast connectivity to VPN sites that span across multiple autonomous systems. This feature was added to MLDP profile that enables Multicast Distribution Trees (MDTs), used for Multicast VPNs (MVPNs), to span multiple autonomous systems.

There are two types of MVPN inter-AS deployment scenarios:

- Single-Provider Inter-AS—A service provider whose internal network consists of multiple autonomous systems.
- Intra-Provider Inter-AS—Multiple service providers that need to coordinate their networks to provide inter-AS support.

To establish a Multicast VPN between two autonomous systems, a MDT-default tunnel must be setup between the two PE routers. The PE routers accomplish this by joining the configured MDT-default group. This MDT-default group is configured on the PE router and is unique for each VPN. The PIM sends the join based on the mode of the groups, which can be PIM SSM, or sparse mode.

Note PIM-Bidir is not supported on MVPN.

Benefits of MVPN Inter-AS Support

The MVPN Inter-AS Support feature provides these benefits to service providers:

- Increased multicast coverage to customers that require multicast to span multiple services providers in an MPLS Layer 3 VPN service.
- The ability to consolidate an existing MVPN service with another MVPN service, as in the case of a company merger or acquisition.

InterAS Option A

InterAS Option A is the basic Multicast VPN configuration option. In this option, the PE router partially plays the Autonomous System Border Router (ASBR) role in each Autonomous System (AS). Such a PE router in each AS is directly connected through multiple VRF bearing subinterfaces. MPLS label distribution protocol need not run between these InterAS peering PE routers. However, an IGP or BGP protocol can be used for route distribution under the VRF.

The Option A model assumes direct connectivity between PE routers of different autonomous systems. The PE routers are attached by multiple physical or logical interfaces, each of which is associated with a given VPN (through a VRF instance). Each PE router, therefore, treats the adjacent PE router like a customer edge (CE) router. The standard Layer 3 MPLS VPN mechanisms are used for route redistribution with each autonomous system; that is, the PEs use exterior BGP (eBGP) to distribute unlabeled IPv4 addresses to each other.

Note Option A allows service providers to isolate each autonomous system from the other. This provides better control over routing exchanges and security between the two networks. However, Option A is considered the least scalable of all the inter-AS connectivity options.

BGP Requirements

PE routers are the only routers that need to be MVPN-aware and able to signal remote PEs with information regarding the MVPN. It is fundamental that all PE routers have a BGP relationship with each other, either directly or through a route reflector, because the PE routers use the BGP peering address information to derive the RPF PE peer within a given VRF.

PIM-SSM MDT tunnels cannot be set up without a configured BGP MDT address-family, because you establish the tunnels, using the BGP connector attribute.

Note A router being RR and PE at that same time for BGP mVPN implementation is not supported, a type 7 and type 6 IPv4 mVPN route is not advertised by a RR, which is also a PE router, if the PE router has the VRF locally configured and when there is a local receiver.

Use full mesh for iBGP mVPN address-family or elect any core (P) router to be the RR.

See the Implementing BGP on Cisco IOS XR Software module of the *Routing Configuration Guide for Cisco NCS 6000 Series Routers* for information on BGP support for Multicast VPN.

Multicast and MVPNv4 over v4GRE Interfaces

Different types of networks rely on the third party network security to attain a secure IP multicast service, which encrypts and decrypts IP unicast traffic across untrusted core network through point-to-point tunnel. Therefore, the customer multicast traffic must be delivered as unicast traffic with encryption across untrusted core network. This is obtained by using generic routing encapsulation (GRE) tunneling to deliver multicast traffic as unicast through tunnel interfaces. Both Multicast and MVPN-v4 over GRE is supported.

- Multicast over v4-GRE Interfaces: Customer networks which are transporting Native IP Multicast across un-trusted core via IPv4 unicast GRE tunnels and encryption.
- MVPN-v4 over GRE Interfaces: Customer networks which are transporting L3VPN multicast services (mVPN-GRE) across an un-trusted core via IPv4 unicast GRE tunnels and encryption.

Note IPv6 Multicast and MVPNv6 over GRE are not supported.

Multicast interface features for GRE tunnels are applied when the inner packet is forwarding through multicast forwarding chain. However, the unicast interface features for GRE underlying interface are applied when the outer transport packet is forwarding through unicast forwarding chain. Thus, multicast interface features such as boundary ACL and TTL threshold are applicable and supported for unicast GRE tunnel just as other multicast main or sub interfaces. However, QoS for unicast GRE tunnel are applied at its underlying physical interface instead of applied on tunnel interface itself.

After setting up unicast routing protocol, the unicast GRE tunnels are treated as interfaces similar to that of a main or sub interface. The unicast GRE tunnels can participate in multicast routing when these are added to multicast routing protocols as multicast enabled interfaces. The unicast GRE tunnels are also used as the accepting or the forwarding interfaces of a multicast route.

Note Traffic drop is observed during In-Service Software Upgrade (ISSU) process for Multicast VPN over generic routing encapsulation (GRE) tunnel.

Concatenation of Unicast GRE Tunnels for Multicast Traffic

This concatenation of unicast GRE tunnels refers to connecting trusted network islands by terminating one unicast GRE tunnel and relaying multicast forwarding to olist that includes different unicast GRE tunnels.

TTL Threshold

GRE enables to workaround networks containing protocols that have limited hop counts. Multicast traffic of mVPN-GRE from encapsulation provider edge (PE) router to decapsulation PE router is considered one hop, and customer packet TTL should be decremented by one number, irrespective of mid-point P routers between these PE routers.

The TTL on GRE transport header is derived from the configuration of GRE tunnel interface, and is decremented when traffic travels from encapsulation PE to decapsulation PE router via P routers. However, for concatenated unicast GRE tunnels, TTL on GRE transport header is reset when the router terminates one unicast GRE tunnel and forwards multicast packet to another unicast GRE tunnel.

Note GRE keep-alive message and the frequency of keep-alive message generation is1 pps. Static police rate in a line card remain 1000 pps to accommodate max 500 unicast GRE tunnel. However, the GRE key is not supported.

Multitopology Routing

Multitopology routing allows you to manipulate network traffic flow when desirable (for example, to broadcast duplicate video streams) to flow over non-overlapping paths.

PIM uses a routing policy that supports matching on source or group address to select the topology in which to look up the reverse-path forwarding (RPF) path to the source. If you do not configure a policy, the existing behavior (to select a default table) remains in force.

Currently, IS-IS and PIM routing protocols alone support multitopology-enabled network.

Label Switched Multicast (LSM) Multicast Label Distribution Protocol (mLDP) based Multicast VPN (mVPN) Support

Label Switch Multicast (LSM) is MPLS technology extensions to support multicast using label encapsulation. Next-generation MVPN is based on Multicast Label Distribution Protocol (mLDP), which can be used to build P2MP and MP2MP LSPs through a MPLS network. These LSPs can be used for transporting both IPv4 and IPv6 multicast packets, either in the global table or VPN context.

Benefits of LSM MLDP based MVPN

LSM provides these benefits when compared to GRE core tunnels that are currently used to transport customer traffic in the core:

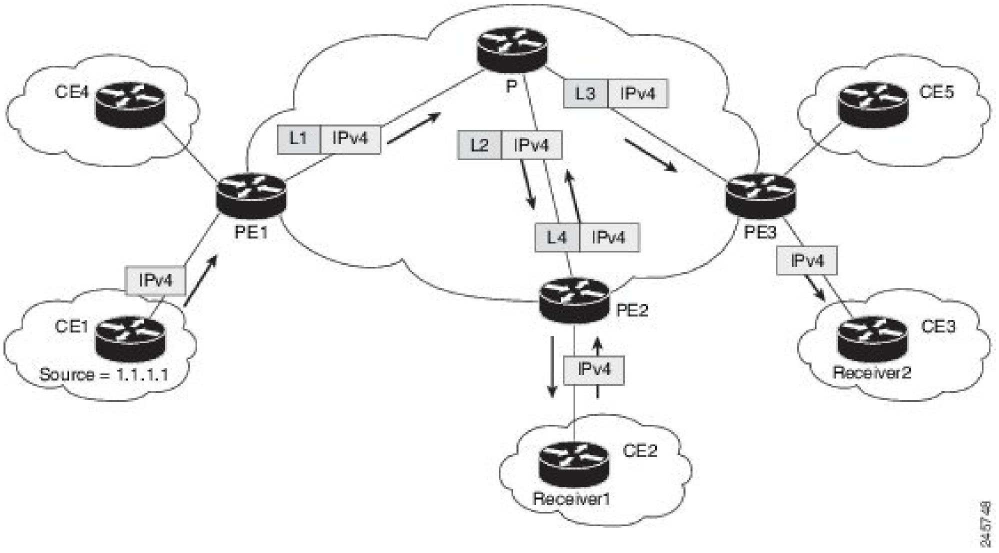
- It leverages the MPLS infrastructure for transporting IP multicast packets, providing a common data plane for unicast and multicast.
- It applies the benefits of MPLS to IP multicast such as Fast ReRoute (FRR) and
- It eliminates the complexity associated PIM.

Configuring MLDP MVPN

The MLDP MVPN configuration enables IPv4 multicast packet delivery using MPLS. This configuration uses MPLS labels to construct default and data Multicast Distribution Trees (MDTs). The MPLS replication is used as a forwarding mechanism in the core network. For MLDP MVPN configuration to work, ensure that the global MPLS MLDP configuration is enabled. To configure MVPN extranet support, configure the source multicast VPN Routing and Forwarding (mVRF) on the receiver Provider Edge (PE) router or configure the receiver mVRF on the source PE. MLDP MVPN is supported for both intranet and extranet.

Note When a is positioned as terminal node, it drops the IPv6 traffic that it receives from Cisco ASR 9000 Series Routers which is acting as a head node since EXP NULL label is sent at the Bottom of stack (BOS) over MLDP tunnel for IPV6 traffic.

Figure 7: MLDP based MPLS Network



P2MP and MP2MP Label Switched Paths

mLDP is an application that sets up Multipoint Label Switched Paths (MP LSPs) in MPLS networks without requiring multicast routing protocols in the MPLS core. mLDP constructs the P2MP or MP2MP LSPs without interacting with or relying upon any other multicast tree construction protocol. Using LDP extensions for MP LSPs and Unicast IP routing, mLDP can setup MP LSPs. The two types of MP LSPs that can be setup are Point-to-Multipoint (P2MP) and Multipoint-to-Multipoint (MP2MP) type LSPs.

A P2MP LSP allows traffic from a single root (ingress node) to be delivered to a number of leaves (egress nodes), where each P2MP tree is uniquely identified with a 2-tuple (root node address, P2MP LSP identifier). A P2MP LSP consists of a single root node, zero or more transit nodes, and one or more leaf nodes, where typically root and leaf nodes are PEs and transit nodes are P routers. A P2MP LSP setup is receiver-driven and is signaled using mLDP P2MP FEC, where LSP identifier is represented by the MP Opaque Value element. MP Opaque Value carries information that is known to ingress LSRs and Leaf LSRs, but need not be interpreted by transit LSRs. There can be several MP LSPs rooted at a given ingress node, each with its own identifier.

A MP2MP LSP allows traffic from multiple ingress nodes to be delivered to multiple egress nodes, where a MP2MP tree is uniquely identified with a 2-tuple (root node address, MP2MP LSP identifier). For a MP2MP LSP, all egress nodes, except the sending node, receive a packet sent from an ingress node.

A MP2MP LSP is similar to a P2MP LSP, but each leaf node acts as both an ingress and egress node. To build an MP2MP LSP, you can setup a downstream path and an upstream path so that:

- Downstream path is setup just like a normal P2MP LSP
- Upstream path is setup like a P2P LSP towards the upstream router, but inherits the downstream labels from the downstream P2MP LSP.

Packet Flow in mLDP-based Multicast VPN

For each packet coming in, MPLS creates multiple out-labels. Packets from the source network are replicated along the path to the receiver network. The CE1 router sends out the native IP multicast traffic. The Provider Edge1 (PE1) router imposes a label on the incoming multicast packet and replicates the labeled packet towards the MPLS core network. When the packet reaches the core router (P), the packet is replicated with the appropriate labels for the MP2MP default MDT or the P2MP data MDT and transported to all the egress PEs. Once the packet reaches the egress PE , the label is removed and the IP multicast packet is replicated onto the VRF interface.

Realizing a mLDP-based Multicast VPN

There are different ways a Label Switched Path (LSP) built by mLDP can be used depending on the requirement and nature of application such as:

- P2MP LSPs for global table transit Multicast using in-band signaling.
- P2MP/MP2MP LSPs for MVPN based on MI-PMSI or Multidirectional Inclusive Provider Multicast Service Instance (Rosen Draft).
- P2MP/MP2MP LSPs for MVPN based on MS-PMSI or Multidirectional Selective Provider Multicast Service Instance (Partitioned E-LAN).

The router performs the following important functions for the implementation of MLDP:

1. Encapsulating VRF multicast IP packet with GRE/Label and replicating to core interfaces (imposition node).
2. Replicating multicast label packets to different interfaces with different labels (Mid node).
3. Decapsulate and replicate label packets into VRF interfaces (Disposition node).

Characteristics of mLDP Profiles

The characteristics of various mLDP profiles are listed in this section.

Profile 4: MS-PMSI-mLDP-MP2MP with BGP-AD

These are the characteristics of this profile:

- MP2MP mLDP trees are used in the core.
- The multicast traffic can be SM or SSM.
- Extranet, Hub and Spoke, CsC, Customer-RP-discovery (Embedded-RP, AutoRP, and BSR) are supported.
- Inter-AS Options A, B, and C are supported. VRF-Route-Import EC is announced in VPN-IP routes.
- Each PE sends Hellos only on the trees rooted on that PE. With this, in deployment scenarios, where the number of source PEs are much lesser than the total number of PEs in the MVPN, results in a huge reduction of PIM neighbors.

Configuration rules for profiles

Rules for mLDP profiles (profile- 4)

- MVPN must be enabled under bgp, if only profile 2 is configured.
- Support only for static RP for customer RP.

MLDP inband signaling

MLDP Inband signaling allows the core to create (S,G) or (*,G) state without using out-of-band signaling such as BGP or PIM. It is supported in VRF (and in the global context). Both IPv4 and IPv6 multicast groups are supported.

In MLDP Inband signaling, one can configure an ACL range of multicast (S,G). This (S,G) can be transported in MLDP LSP. Each multicast channel (S,G), is 1 to 1 mapped to each tree in the inband tree. The (S,G) join, through IGMP/MLD/PIM, will be registered in MRIB, which is the client of MLDP.

MLDP In-band signalling supports transiting PIM (S,G) or (*,G) trees across a MPLS core without the need for an out-of-band protocol. In-band signaling is only supported for shared-tree-only forwarding (also known as sparse-mode threshold infinity). PIM Sparse-mode behavior is not supported (switching from (*,G) to (S,G).

The details of the MLDP profiles are discussed in the *Multicast Configuration Guide for the Cisco NCS 6000 Series Routers*

Summary of Supported MVPN Profiles

This tables summarizes the supported MVPN profiles:

Profile Number	Name	Opaque-value	BGP-AD	Data-MDT
4	MS- PMSI (Partition) MLDP MP2MP with BGP -AD	Type 1 - Source-PE:Global -ID	• I- PMSI with empty PTA • MS- PMSI for partition mdt • S- PMSI for data-mdt • S- PMSI cust RP-discovery trees	BGP-AD

Multicast Source Discovery Protocol

Multicast Source Discovery Protocol (MSDP) is a mechanism to connect multiple PIM sparse-mode domains. MSDP allows multicast sources for a group to be known to all rendezvous points (RPs) in different domains. Each PIM-SM domain uses its own RPs and need not depend on RPs in other domains.

An RP in a PIM-SM domain has MSDP peering relationships with MSDP-enabled routers in other domains. Each peering relationship occurs over a TCP connection, which is maintained by the underlying routing system.

MSDP speakers exchange messages called Source Active (SA) messages. When an RP learns about a local active source, typically through a PIM register message, the MSDP process encapsulates the register in an SA message and forwards the information to its peers. The message contains the source and group information for the multicast flow, as well as any encapsulated data. If a neighboring RP has local joiners for the multicast group, the RP installs the S, G route, forwards the encapsulated data contained in the SA message, and sends PIM joins back towards the source. This process describes how a multicast path can be built between domains.

Note Although you should configure BGP or Multiprotocol BGP for optimal MSDP interdomain operation, this is not considered necessary in the Cisco IOS XR Software implementation. For information about how BGP or Multiprotocol BGP may be used with MSDP, see the MSDP RPF rules listed in the Multicast Source Discovery Protocol (MSDP), Internet Engineering Task Force (IETF) Internet draft.

Multicast Nonstop Forwarding

The Cisco IOS XR Software nonstop forwarding (NSF) feature for multicast enhances high availability (HA) of multicast packet forwarding. NSF prevents hardware or software failures on the control plane from disrupting the forwarding of existing packet flows through the router.

The contents of the Multicast Forwarding Information Base (MFIB) are frozen during a control plane failure. Subsequently, PIM attempts to recover normal protocol processing and state before the neighboring routers time out the PIM hello neighbor adjacency for the problematic router. This behavior prevents the NSF-capable router from being transferred to neighbors that will otherwise detect the failure through the timed-out adjacency. Routes in MFIB are marked as stale after entering NSF, and traffic continues to be forwarded (based on those routes) until NSF completion. On completion, MRIB notifies MFIB and MFIB performs a mark-and-sweep to synchronize MFIB with the current MRIB route information.

Multicast Configuration Submodes

Cisco IOS XR Software moves control plane CLI configurations to protocol-specific submodes to provide mechanisms for enabling, disabling, and configuring multicast features on a large number of interfaces.

Cisco IOS XR Software allows you to issue most commands available under submodes as one single command string from the global or XR config mode.

For example, the **ssm** command could be executed from the PIM configuration submode like this:

```
RP/0/RSP0/CPU0:router(config)# router pim
RP/0/RSP0/CPU0:router(config-pim)# address-family ipv4
RP/0/RSP0/CPU0:router(config-pim-default-ipv4)# ssm range
```

Alternatively, you could issue the same command from the global or XR config mode like this:

```
RP/0/RSP0/CPU0:router(config)# router pim ssm range
```

The following multicast protocol-specific submodes are available through these configuration submodes:

Multicast-Routing Configuration Submode

When you issue the **multicast-routing ipv4** command, all default multicast components (PIM, IGMP, , MFWD, and MRIB) are automatically started, and the CLI prompt changes to “config-mcast-ipv4” indicating that you have entered multicast-routing configuration submode.

PIM Configuration Submode

When you issue the **router pim** command, the CLI prompt changes to “config-pim-ipv4,” indicating that you have entered the default pim address-family configuration submode.

IGMP Configuration Submode

When you issue the **router igmp** command, the CLI prompt changes to “config-igmp,” indicating that you have entered IGMP configuration submode.

MSDP Configuration Submode

When you issue the **router msdp** command, the CLI prompt changes to “config-msdp,” indicating that you have entered router MSDP configuration submode.

Understanding Interface Configuration Inheritance

Cisco IOS XR Software allows you to configure commands for a large number of interfaces by applying command configuration within a multicast routing submode that could be inherited by all interfaces. To override the inheritance mechanism, you can enter interface configuration submode and explicitly enter a different command parameter.

For example, in the following configuration you could quickly specify (under router PIM configuration mode) that all existing and new PIM interfaces on your router will use the hello interval parameter of 420 seconds. However, Packet-over-SONET/SDH (POS) interface 0/1/0/1 overrides the global interface configuration and uses the hello interval time of 210 seconds.

```
RP/0/RP0/CPU0:router(config)# router pim
RP/0/RP0/CPU0:router(config-pim-default-ipv4)# hello-interval 420
RP/0/RP0/CPU0:router(config-pim-default-ipv4)# interface pos 0/1/0/1
RP/0/RP0/CPU0:router(config-pim-ipv4-if)# hello-interval 210
```

Understanding Interface Configuration Inheritance Disablement

As stated elsewhere, Cisco IOS XR Software allows you to configure multiple interfaces by applying configurations within a multicast routing submode that can be inherited by all interfaces.

To override the inheritance feature on specific interfaces or on all interfaces, you can enter the address-family IPv4 submode of multicast routing configuration mode, and enter the **interface-inheritance disable** command together with the **interface** *type interface-path-id* or **interface all** command. This causes PIM or IGMP protocols to disallow multicast routing and to allow only multicast forwarding on those interfaces specified. However, routing can still be explicitly enabled on specified individual interfaces.

For related information, see Understanding Enabling and Disabling Interfaces, on page 26

Understanding Enabling and Disabling Interfaces

When the Cisco IOS XR Software multicast routing feature is configured on your router, by default, no interfaces are enabled.

To enable multicast routing and protocols on a single interface or multiple interfaces, you must explicitly enable interfaces using the **interface** command in multicast routing configuration mode.

To set up multicast routing on all interfaces, enter the **interface all** command in multicast routing configuration mode. For any interface to be fully enabled for multicast routing, it must be enabled specifically (or be default) in multicast routing configuration mode, and it must not be disabled in the PIM and IGMP configuration modes.

For example, in the following configuration, all interfaces are explicitly configured from multicast routing configuration submode:

```
RP/0/RP0/CPU0:router(config)# multicast-routing
RP/0/RP0/CPU0:router(config-mcast)# interface all enable
```

To disable an interface that was globally configured from the multicast routing configuration submode, enter interface configuration submode, as illustrated in the following example:

```
RP/0/RP0/CPU0:router(config-mcast)# interface GigabitEthernet0pos 0/1/0/0
RP/0/RP0/CPU0:router(config-mcast-default-ipv4-if)# disable
```

Multicast Routing Information Base

The Multicast Routing Information Base (MRIB) is a protocol-independent multicast routing table that describes a logical network in which one or more multicast routing protocols are running. The tables contain generic multicast routes installed by individual multicast routing protocols. There is an MRIB for every logical network in which the router is configured. MRIBs do not redistribute routes among multicast routing protocols; they select the preferred multicast route from comparable ones, and they notify their clients of changes in selected attributes of any multicast route.

Multicast Forwarding Information Base

Multicast Forwarding Information Base (MFIB) is a protocol-independent multicast forwarding system that contains unique multicast forwarding entries for each source or group pair known in a given network. There is a separate MFIB for every logical network in which the router is configured. Each MFIB entry resolves a given source or group pair to an incoming interface (IIF) for reverse forwarding (RPF) checking and an outgoing interface list (olist) for multicast forwarding.

MSDP MD5 Password Authentication

MSDP MD5 password authentication is an enhancement to support Message Digest 5 (MD5) signature protection on a TCP connection between two Multicast Source Discovery Protocol (MSDP) peers. This feature provides added security by protecting MSDP against the threat of spoofed TCP segments being introduced into the TCP connection stream.

MSDP MD5 password authentication verifies each segment sent on the TCP connection between MSDP peers. The **password clear** command is used to enable MD5 authentication for TCP connections between two MSDP peers. When MD5 authentication is enabled between two MSDP peers, each segment sent on the TCP connection between the peers is verified.

Note MSDP MD5 authentication must be configured with the same password on both MSDP peers to enable the connection between them. The 'password encrypted' command is used only for applying the stored running configuration. Once you configure the MSDP MD5 authentication, you can restore the configuration using this command.

MSDP MD5 password authentication uses an industry-standard MD5 algorithm for improved reliability and security.

Multicast VPN IPv6 over IPv4 GRE

MVPN Overview

Multicast VPN is based on the multicast domain, in which PE routers establish virtual PIM neighbor connections with other PE routers connected to the same customer VPN, and form a secure, virtual multicast domain over the provider network. Customer multicast traffic is transmitted across the core from one site to another as if the traffic flows through a dedicated provider network.

Secure data transmission is achieved by encapsulating customer multicast packets in a provider header using GRE encapsulation at Ingress PE routers and then transmitting the new provider packets across the core. At Egress PE routers, packets are decapsulated to reveal the original customer packet and they are sent to CE routers downstream. The encapsulation and decapsulation operations and the packet transmission inside the core are abstracted by the use of a new virtual tunnel interface called MDT (Multicast Distribution Tree) tunnel.

To separate one VPN's traffic from another VPN's traffic, dedicated multicast routing and forwarding tables are created for each VPN. VPN-specific multicast routing and forwarding database is referred to as MVRF. On a PE router, when multicast is enabled for a VRF, a MVRF is created. Multicast routing protocols like PIM, IGMP, and MSDP run in the context of a MVRF. Similarly, all routes created by a MVRF protocol instance are associated with the corresponding MVRF. In addition to MVRFs, which holds VPN specific protocol states, a PE router always has a global VRF, which contains all the routing and forwarding information for the provider network.

Encapsulating customer packets in a provider header makes P routers unaware of the origin of the packets. Instead, all VPN packets passing through the provider network are viewed as native multicast packets and are routed according to the routing information in the core network. P routers do not keep any customer routing information. To support MVPN, they only need to support native multicast routing. PE routers, on the other hand, are upgraded to handle MVPN. Avoiding changes to P routers significantly reduces the risk associated with MVPN deployment, and therefore guarantees the stability of the provider network.

Due to the separation of customer and provider routing information, customers and providers are free to use their own multicast configurations, and choose different PIM modes, such as SM, BIDIR, and SSM.

Note that any packet sent to the default MDT tunnel is forwarded to all the PE routers that are part of the customer VRF, irrespective of whether they are interested in the particular customer VRF route, or not.

MVPN IPv6 over IPv4 GRE

MVPN IPv6 traffic can be carried over IPv4 GRE multicast tunnels using the standard GRE tunneling technique that is designed to provide the services to implement any standard point-to-point encapsulation scheme. As in IPv6 manually configured tunnels, GRE tunnels are links between two points, with a separate tunnel for each link. The tunnels are not tied to a specific passenger or transport protocol but, in this case, carry IPv6 as the passenger protocol with the GRE as the carrier protocol and IPv4 or IPv6 as the transport protocol.

The primary use of GRE tunnels is for stable connections that require regular secure communication between two edge devices or between an edge device and an end system.

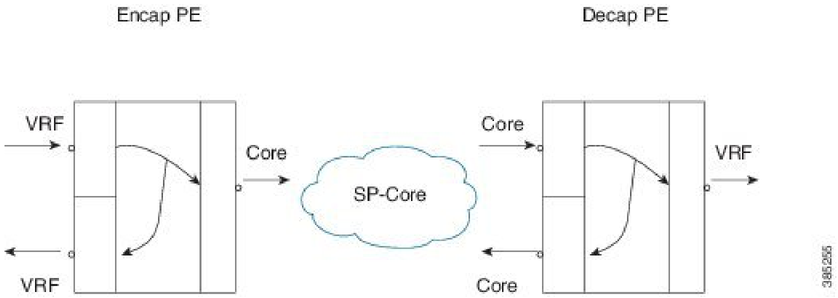
MVPN IPv6 over IPv4 GRE Support

Protocol Independent Multicast in Sparse Mode (PIM-SM) supports this feature at the customer side and Protocol Independent Multicast in Source-Specific Multicast (PIM-SSM) supports this feature at the core. This feature is supported only on Profile 0 (Rosen GRE without BGP AD).

Traffic Forwarding Behavior

This section describes the traffic forwarding scheme in terms of operations of the Encap (CE to PE) and Decap (PE to CE) directions for ingress and egress LCs.

Figure 8: Forwarding Scheme



In Figure 1, the VRF interfaces are customer facing. The Core interfaces face the SP Core. A customer packet enters the Encap PE router and gets forwarded to the core, in addition to another VRF interface (local switching). Similarly, on the Decap PE router, the packet from the core gets forwarded to the VRF interface, in addition to another core interface (the only path to reach another Core P/PE router is through this PE).

Encapsulation PE

Ingress (VRF interface)

1. From the interface UIDB, the ucode knows whether the packet arrived on a VRF interface and also knows the VRF table ID.
2. Route lookup takes place in the VRF table (s1, g1, VRF table ID).
3. Packet matches VRF route (s1, g1).
4. The route lookup provides the following information:
   - Flag to indicate whether the packet needs encapsulation
   - Encap ID
   - FGID
5. Packet is forwarded to the fabric with encap ID and FGID.

Egress (Core / VRF interfaces)

On egress LC:

- Receive packets from fabric after replication.
- Encap ID lookup takes place. The Encap-ID result has pointers to each encapsulation chain.
- Encapsulates the packet and sends over the core.
- If there are local receivers on this router, Encap-ID result indicates the same. Forwards the packet natively.

Decapsulation PE

Ingress (Core)

1. The interface UIDB indicates that it is not a VRF interface.
2. Route lookup takes place in the global table, packet matches global route (S,G).
3. RPF check is done using the physical input interface of the packet.
4. The core route lookup yields:
    - A flag indicating whether this is a tunnel route (tunnel flag)
    - A flag indicating whether the outer header should be stripped (decap flag). It is used when the core route does not have any core interfaces in the olist.
    - Tunnel UIDB
5. If the decap flag is set, packet is decapsulated and forwarded to fabric with tunnel UIDB.

Egress (Core / VRF interfaces)

The egress processing is identical to the Encap PE case described earlier, except that after decapsulating the packet, the following additional steps are taken:

- The incoming interface is set to the tunnel interface.
- The RPF check is done to ensure that we can accept the packet from the tunnel interface.
- Lookup takes place on VRF table based on the tunnel UIDB and forwards the packet to the customer.

MVPN IPv6 over IPv4 GRE Configuration: Examples

These examples show how to configure MVPN IPv6 over IPv4 GRE:

Multicast-routing

```
multicast-routing
  address-family ipv4
   mdt source Loopback0
   maximum disable
   rate-per-route
   interface all enable
   accounting per-prefix
  !
```

```
vrf blue
   address-family ipv4
    mdt data 238.1.1.1/32
    mdt default ipv4 239.1.1.1
    mdt data 255 threshold 2
    rate-per-route
    interface all enable
    accounting per-prefix
   !
  !
 !
```

PIM:

```
router pim
  address-family ipv4
   rp-address 1.1.1.1
  !
  vrf blue
   address-family ipv4
    rp-address 2.2.2.2
   !
  !
 !
```

Point-to-Multipoint Traffic Engineering with GTM

Point-to-Multipoint Traffic Engineering Overview

The Point-to-Multipoint (P2MP) Resource Reservation Protocol-Traffic Engineering (RSVP-TE) solution allows service providers to implement IP multicast applications, such as IPTV and real-time video, broadcast over the MPLS label switch network. The RSVP-TE protocol is extended to signal point-to-point (P2P) and P2MP label switched paths (LSPs) across the MPLS networks. By using RSVP-TE extensions as defined in RFC 4875, multiple sub-LSPs are signaled for a given TE source.

Service providers use native IP multicast in both core and access network for providing services like IPTV or content delivery. This requires running a multicast routing protocol in both their core network and their access network.

With the proliferation of MPLS in IP network, it is increasingly popular for ISP to leverage MPLS technology for transporting IP multicast traffic. One popular demand is to use TE-FRR for transporting mission critical multicast data over the network. In the event of link or node failure, FRR allows traffic to quickly recover from network failures and thus minimizes any traffic disruption.

Most of the label distribution mechanisms (RSVP-TE, LDP and BGP) primarily focus on signaling P2P (Point-to-Point) LSP. Because there is always a single source and destination for each flow, the P2P LSP is sufficient for IP unicast. In multicast, however, the source is sending traffic to an arbitrary group of hosts. In fact, multicast data distribution is often represented as a "tree" where the source is the root and receivers are "leaf" of the tree.

For optimal traffic utilization, the IP/MPLS core routers perform label replication to support the transport of multicast traffic. In such cases, the LSP is described as P2MP. The P2MP LSP is capable of determining optimal branch points in the IP/MPLS network and therefore offers better link utilization.

The following technologies are supported to apply MPLS for multicast service:

- Ingress Node makes a mapping between multicast traffic and a LSP in order to forward the multicast traffic to the LSP.
- RSVP-TE signals and establishes a P2MP LSP in order to forward the multicast traffic over the LSP.
- MPLS forwarding scheme supports the labeled packet replication similar to that of the the IP multicast forwarding scheme.
- Egress Node conducts a special RPF check because the multicast traffic is encapsulated with MPLS and no multicast routing protocol runs in the core network.

P2MP TE with GTM

P2MP TE with GTM (Global Table Multicast) is available only in the Global table. The (S,G) route in the global table can be mapped to P2MP TE tunnels. However, this feature now enables service providers to use P2MP TE tunnels to carry VRF multicast traffic. Static mapping is used to map VRF (S, G) traffic to P2MP TE tunnels, and BGP-AD is used to send P2MP BGP opaque that includes VRF-based P2MP FEC as MDT Selective Provider Multicast Service Interface (S-PMSI).

Note For P2MP GTM profile, only tail node (egress) functionality is supported. NCS6k as only decap node (egress) is supported for P2MP GTM profile.

Global Table Multicast and Egress PE

Multicast routing information that is not specific to VPNs is stored in a router's "global table", rather than in a VRF; therefore, it is known as "Global Table Multicast" (GTM).

In the case where global table multicast uses PIM-SM in ASM (Any Source Multicast) mode, the inter-area P2MP service LSP can be used to carry traffic either on a shared (*,G) or a source (S,G) tree.

An egress PE learns the (S/*,G) of a multicast stream as a result of receiving IGMP or PIM messages on one of its IP multicast interfaces. This (S/*,G) forms the P2MP FEC of the inter-area P2MP service LSP. For each of such a P2MP FEC, a distinct inter-area P2MP service LSP may exist, or multiple FECs may be carried over a single P2MP service LSP using a wildcard (*,*) S-PMSI.

For P2MP TE GTM profile, only tail node (egress) functionality is supported in Release 6.1.1. Only static RP-RP selection type is supported for P2MP TE GTM profile. The following are not supported:

- Bud node scenario
- Auto RP - RP selection type
- P2MP TE GTM profile for IPv6

P2MP TE Functional Overview

Tunnel Head

At the Headend, RSVP-TE signals and establishes a P2MP TE LSP. In order to forward the multicast traffic over the P2MP TE LSP, the multipoint tunnel interface is enabled for multicast forwarding. In addition, IGMPv2/IGMPv3 or MLDv2 is statically configured to forward multicast streams on the P2MP tunnel.

Additionally, MPLS and RSVP-TE are activated on each MPLS core-facing interface.

Midpoint

At mid node, no multicast routing protocol runs. Therefore, there is no configuration related to the multicast routing. However, a global configuration is required for activating LMRIB function in the multicast package. This configuration is set at all nodes that process P2MP LSPs.

MPLS and RSVP-TE are activated on each MPLS core-facing interface.

The mid-node signals and establishes a P2MP TE LSP according to the signaling message from the upstream node. This node conducts the label packet replication for forwarding the packet to each downstream node.

Tunnel Tail

At the Tailend, RSVP-TE signals and establishes a P2MP TE LSP. MPLS and RSVP-TE are thus configured on each MPLS core-facing interface.

PIM and PIMv6 conduct a special RPF check for the packet that is received from MPLS core.

Instead of the normal IP RPF check, the tail end node checks whether the packet comes from the correct headend, or not. If not, this node drops the received packet.

P2MP TE Forwarding Methodology

P2MP TE tunnel based forwarding uses a two-stage forwarding model. On the P2MP tunnel head, on the ingress line card, IP multicast packets are encapsulated as MPLS using a local label corresponding to the P2MP tunnel. These packets are multicast over the fabric to egress line cards hosting the P2MP tunnel. At the tunnel midpoint and tail, on ingress line card, packets are received as MPLS and sent to egress LC as MPLS. Therefore, on P2MP tunnel head, midpoint and tailend, multicast packets are received as MPLS on the egress LC. Egress LC processes these packets and replicates them on each outgoing leg of the P2MP tunnel by swapping the top label with the outgoing label for each leg. The egress LC can also pop the label and send the packet to MFIB for forwarding the packet as pure IP on a bud node or tunnel tail along with the RPF information. The RPF check is performed by the MFIB when forwarding the packet as pure IP.

Egress PE

On the Egress PE, incoming packet on P2MP tunnel is MPLS. On the ingress LC, the MPLS table lookup that is based on the top MPLS label gives FGIDs (Primary and Backup FGIDs) and FRR (Fast Reroute)active information. The MPLS packet is sent over the fabric using the FGID if FRR is not active on the tunnel. If FRR is active on the tunnel, the packet is sent over the fabric using the backup FGID.

The processing on the egress LC on the egress PE is exactly same as the processing on the egress LC on the ingress PE.

For P2MP TE GTM profile, only tail node (egress) functionality is supported in Release 6.1.2.

Fast Reroute

Fast Reroute (FRR) is a mechanism to minimize interruptions in traffic delivery, as a result of link failures, to a TE Label Switched Path (LSP) destination. FRR enables temporarily fast switching of LSP traffic along an alternative backup path around a network failure, until the TE tunnel source signals a new end-to-end LSP.

To facilitate FRR, the FIB process on each node gets informed of the protected interface going down. The high priority FIB FRR thread receives the notification and takes FRR action to quickly reroute the traffic to the backup tunnel, within a short duration of time (in the order of ms). Then the FIB FRR thread pulses the FIB main thread which performs FRR recovery actions. Finally, RSVP moves the sub-LSPs over to the backup-tunnel.

Tunnel Reoptimization

When a router looks to determined whether there is a better path for tunnels that are already up, the process is known as tunnel reoptimization. P2MP tunnel reoptimization takes place at the tunnel headend. During tunnel reoptimization the whole LSP tree is recomputed and signaled.

P2MP TE with GTM Configuration

This section shows how to configure Point-to-Multipoint TE with GTM.

```
router bgp 65488
 address-family ipv4 mvpn
  global-table-multicast
!
af-group IPV4_MULTICAST-GTM address-family ipv4 mvpn
  next-hop-self
!
neighbor-group BBR-PEERS
   !
  address-family ipv4 mvpn
   use af-group IPV4_MULTICAST-GTM



route-policy GTM
  set core-tree p2mp-te-default
end-policy
!
router pim
address-family ipv4
  rpf topology route-policy GTM
  mdt c-multicast-routing bgp
  !
  rp-address 100.1.2.1
  maximum routes 200000

  !
!
!

multicast-routing
address-family ipv4
  interface all enable
# uplinks and crosslinks disabled
  interface POS0/6/0/1
   disable
  !
  interface GigabitEthernet0/6/1/4
   disable
  !
  interface GigabitEthernet0/6/1/5
   disable
  !
  interface GigabitEthernet0/6/2/0
   disable
  !
  mdt source Loopback0
  export-rt 701:0
  import-rt 701:0
  !
  bgp auto-discovery p2mp-te
```

```
    receiver-site
   !
   mdt default p2mp-te
   !
   #monitoring for "show mfib route rate" and "sh mfib route statistics 232.0.0.10 loc
0/6/cpu0"
   rate-per-route
   accounting per-prefix
 !
 !

mpls traffic-eng
auto-tunnel p2mp
  tunnel-id min 2100 max 2200
```

Multicast Label Distribution Protocol Edge

The Multicast Label Distribution Protocol (MLDP) feature is enhanced to support the edges; that is, the encapsulation (headend) and the decapsulation (tailend) at the Provider Edge (PE) devices.

This feature enables service providers to extend the existing MPLS backbone network for multicast services. It extends the functionality from midpoint to support the edges: the headend and the tailend.

Earlier, Profile 4- MS-PMSI MLDP MP2MP (Multipoint-to-Multipoint) with BGP AD was supported at the core, now it is supported on the edge.

The following characteristics support Profile 4:

- MP2MP MLDP trees
- BGP AD enabled
- Customer traffic may be SM or SSM
- IPv4 and IPv6 supported

MLDP Edge Configuration: Example

This example shows how to configure MDT MLDP MP2MP MVPN with BGP-AD:

```
router bgp 100
mvpn
bgp router-id 100.0.0.1
 bgp graceful-restart
 address-family ipv4 unicast
address-family vpnv4 unicast
address-family ipv6 unicast
address-family vpnv6 unicast
address-family ipv4 mvpn
address-family ipv6 mvpn
 !
 neighbor 100.0.0.3
  remote-as 100
  update-source Loopback0
  address-family ipv4 unicast
   next-hop-self
  !
  address-family vpnv4 unicast
  !
  address-family vpnv6 unicast
!
  address-family ipv4 mvpn
 !
  address-family ipv6 mvpn
!
 neighbor 100.0.0.4
  remote-as 100
  update-source Loopback0
  address-family ipv4 unicast
   next-hop-self
  !
  address-family vpnv4 unicast
  !
  address-family vpnv6 unicast
!
  address-family ipv4 mvpn
 !
  address-family ipv6 mvpn
  !
 vrf p4_v46
  rd 104:1
  address-family ipv4 unicast
  address-family ipv6 unicast
  address-family ipv4 mvpn
  address-family ipv6 mvpn
  neighbor 15.0.0.2
   remote-as 200
   address-family ipv4 unicast
    route-policy PASS in
    route-policy PASS out
    as-override
  !
  neighbor 2008:15::2
   remote-as 200
   address-family ipv6 unicast
    route-policy PASS in
    route-policy PASS out
    as-override
!
!
mpls ldp
 router-id 100.0.0.1
 mldp
  make-before-break delay 30 0
 !
 interface TenGigE0/0/0/0
 !
!
multicast-routing
 address-family ipv4
  mdt source Loopback0
  rate-per-route
  interface all enable
  accounting per-prefix
 !
 address-family ipv6
  rate-per-route
  interface all enable
  accounting per-prefix
 !
 vrf p4_v46
  address-family ipv4
   bgp auto-discovery mldp
   mdt partitioned mldp ipv4 mp2mp
   rate-per-route
   interface all enable
   accounting per-prefix
  !
  address-family ipv6
   bgp auto-discovery mldp
   mdt partitioned mldp ipv4 mp2mp
   rate-per-route
   interface all enable
   accounting per-prefix
!
!
router pim
 vrf p4_v46
  address-family ipv4
   rpf topology route-policy partition-mp2mp
  !
  address-family ipv6
   rpf topology route-policy partition-mp2mp
!
!
end


route-policy partition-mp2mp
 set core-tree mldp-partitioned-mp2mp
end-policy
```

Verifying MLDP Edge Control Plane and Trouble Shooting: Examples

The following sequence of examples show how to verify MLDP Edge control plane and troubleshoot on a router:

```
 show pim vrf vrf_104 ipv4 neighbor

Thu Oct  8 15:08:42.077
PIM neighbors in VRF vrf_104
 Flag: B - Bidir capable, P - Proxy capable, DR - Designated Router,      E - ECMP Redirect
 capable       * indicates the neighbor created for this router

 Neighbor Address             Interface              Uptime    Expires  DR pri   Flags
 10.1.104.1*                  TenGigE0/1/0/4/4.104   00:02:28  00:01:27 1 (DR) B
 10.1.104.2                   TenGigE0/1/0/4/4.104   00:02:27  00:01:28 0
 100.1.255.254*               Lmdtvrf/104            01:22:15  00:01:41 1 (DR)
100.1.255.104*               Loopback104            01:22:15  00:01:22 1 (DR) B E
```

```
show mrib vrf vrf_104 ipv4 route 232.1.2.4

Thu Oct  8 15:08:42.666 UTC

 IP Multicast Routing Information Base Entry flags: L - Domain-Local Source, E - External
Source to the Domain,     C - Directly-Connected Check, S - Signal, IA - Inherit Accept,
   IF - Inherit From, D - Drop, ME - MDT Encap, EID - Encap ID,     MD - MDT Decap, MT -
MDT Threshold Crossed, MH - MDT interface handle     CD - Conditional Decap, MPLS - MPLS
Decap, EX - Extranet     MoFE - MoFRR Enabled, MoFS - MoFRR State, MoFP - MoFRR Primary
  MoFB - MoFRR Backup, RPFID - RPF ID Set, X - VXLAN  Interface flags: F - Forward, A -
Accept, IC - Internal Copy,    NS - Negate Signal, DP - Don't Preserve, SP - Signal Present,
     II - Internal Interest, ID - Internal Disinterest, LI - Local Interest,     LD - Local
 Disinterest, DI - Decapsulation Interface     EI - Encapsulation Interface, MI - MDT
Interface, LVIF - MPLS Encap,     EX - Extranet, A2 - Secondary Accept, MT - MDT Threshold
 Crossed,     MA - Data MDT Assigned, LMI - mLDP MDT Interface, TMI - P2MP-TE MDT Interface
     IRMI - IR MDT Interface

 (10.1.104.2,232.1.2.4) RPF nbr: 10.1.104.2 Flags: RPF MT^M
  MT Slot: 5/2/4
  Up: 00:02:27
 Incoming Interface List
     TenGigE0/1/0/4/4.104 Flags: A, Up: 00:02:27
   Outgoing Interface List
    Lmdtvrf/104 Flags: F LMI MT MA TR, Up: 00:02:27
```

show mfib vrf vrf_104 ipv4 route 232.1.2.4 detail

```
Thu Oct  8 15:08:43.160 UTC

IP Multicast Forwarding Information Base
Entry flags: C - Directly-Connected Check, S - Signal, D - Drop,   IA - Inherit Accept, IF
 - Inherit From, EID - Encap IDM   ME - MDT Encap, MD - MDT Decap, MT - MDT Threshold
Crossed,   MH - MDT interface handle, CD - Conditional Decap,   DT - MDT Decap True, EX -
Extranet, RPFID - RPF ID Set,
MoFE - MoFRR Enabled, MoFS - MoFRR State, X - VXLAN Interface flags: F - Forward, A - Accept,
 IC - Internal Copy,   NS - Negate
Signal, DP - Don't Preserve, SP - Signal Present,   EG - Egress, EI - Encapsulation Interface,
 MI - MDT Interface,   EX - Extranet, A2 - Secondary Accept
 Forwarding/Replication Counts: Packets in/Packets out/Bytes out Failure Counts: RPF / TTL
 / Empty Olist / Encap RL / Other

 (10.1.104.2,232.1.2.4),   Flags:  EID , FGID: 29087, -1, 29088 ,
   Up: 00:02:28
   Last Used: never
   SW Forwarding Counts: 0/0/0
   SW Replication Counts: 0/0/0
  SW Failure Counts: 0/0/0/0/0
   Route ver: 0xad86
  MVPN Info :-
Associated Table ID : 0xe0000000
    MDT Handle: 0x0, MDT Probe:N [N], Rate:Y, Acc:Y
    MDT SW Ingress Encap V4/V6, Egress decap: 0 / 0, 0
     Encap ID: 7, RPF ID: 0
    Local Receiver: False, Turnaround: False
    In AMT List: False
   Lmdtvrf/104 Flags:  F LMI TR, Up:00:02:28 (stale, mt, )
   TenGigE0/1/0/4/4.104 Flags:  A, Up:00:02:28 (stale, mt, )
RP/0/RP0/CPU0:NCS-MLDP-R1#
```

show pim vrf vrf_104 ipv4 group-map

```
Thu Oct  8 15:08:43.680 UTC

 IP PIM Group Mapping Table
(* indicates group mappings being used)
(+ indicates BSR group mappings active in MRIB)
 Group Range         Proto Client   Groups RP address      Info  224.0.1.39/32*      DM
 perm     0      0.0.0.0          224.0.1.40/32*      DM    perm     1      0.0.0.0
   224.0.0.0/24*       NO    perm     0      0.0.0.0          232.0.0.0/8*        SSM
config   100    0.0.0.0          224.0.0.0/4*        SM    config   0      100.1.255.104
 RPF: De3,100.1.255.104 (us)
224.0.0.0/4         SM    static   0      0.0.0.0         RPF: Null,0.0.0.0
```

show bgp vrf vrf_104 ipv4 mvpn

```
Thu Oct  8 15:08:44.456 UTC
BGP VRF vrf_104, state: Active
 BGP Route Distinguisher: 104:1 VRF ID: 0x60000006
 BGP router identifier 100.1.255.254, local AS number 200
 Non-stop routing is enabled
BGP table state: Active
 Table ID: 0x0   RD version: 521
 BGP main routing table version 521
 BGP NSR Initial initsync version 14 (Reached)
 BGP NSR/ISSU Sync-Group versions 0/0 Status codes: s suppressed, d damped, h history, *
valid, > best               i - internal, r RIB-failure, S stale, N Nexthop-discard Origin
 codes: i - IGP, e - EGP, ? - incomplete
 Network            Next Hop            Metric LocPrf Weight Path
 Route Distinguisher: 104:1 (default for vrf vrf_104)
*> [1][100.1.255.254]/40
                      0.0.0.0                                0 i
 *>i[1][100.2.255.254]/40
                       100.2.255.254                 100      0 i
 *> [3][0][0.0.0.0][0][0.0.0.0][100.1.255.254]/120
                     0.0.0.0                                0 i
*>i[3][0][0.0.0.0][0][0.0.0.0][100.2.255.254]/120
                     100.2.255.254                 100      0 i
*> [3][32][10.1.104.2][32][232.1.2.4][100.1.255.254]/120
                       0.0.0.0                                0 i
 *> [3][32][10.1.104.2][32][232.1.2.5][100.1.255.254]/120
                       0.0.0.0                                0 i
```

show mpls mldp neighbors

```
Thu Oct  8 15:08:45.607 UTC
mLDP neighbor database
 MLDP peer ID      : 100.3.255.254:0, uptime 01:17:33 Up,   Capabilities     : GR, Typed
Wildcard FEC, P2MP, MP2MP, MBB
  Target Adj       : No
   Upstream count   : 0
   Branch count     : 305
   LDP GR           : Enabled
                    : Instance: 1
 Label map timer  : never
  Policy filter in :
  Path count       : 4
   Path(s)          : 31.3.0.1          TenGigE0/0/0/0/9 LDP

                   : 31.6.0.1          TenGigE0/1/0/4/1 LDP
                  : 31.7.0.1          TenGigE0/1/0/4/2 LDP
                    : 31.8.0.1          TenGigE0/1/0/9/2 LDP
   Adj list         : 31.3.0.1          TenGigE0/0/0/0/9
                    : 31.6.0.1          TenGigE0/1/0/4/1
                    : 31.7.0.1          TenGigE0/1/0/4/2
                    : 31.8.0.1          TenGigE0/1/0/9/2
   Peer addr list   : 100.3.255.254
                    : 4.4.2.12
                    : 30.1.0.1
                    : 32.2.0.1
                    : 32.3.0.1
                    : 32.4.0.1
                    : 32.1.0.1
                    : 31.1.0.1
                    : 31.3.0.1
                    : 31.6.0.1
                    : 31.7.0.1
                    : 31.8.0.1
```

show mrib vrf vrf_104 ipv4 route summary

```
 Thu Oct  8 15:08:46.116 UTC
 MRIB Route Summary for VRF vrf_104
  No. of group ranges = 5
   No. of (*,G) routes = 1
   No. of (S,G) routes = 100
   No. of Route x Interfaces (RxI) = 100
   Total No. of Interfaces in all routes = 202
```

 show mvpn vrf vrf_104 ipv4 context detail | in HLI

```
Thu Oct  8 15:08:46.592 UTC
  MLDP Number of Roots: 0 (Local: 0), HLI: 0x00000, Rem HLI: 0x00000   Partitioned MDT:
Configured, MP2MP (RD:Not added, ID:Added), HLI: 0x00007, Loc Label: 24002, Remote: Configured
     ID: 3 (0x15587a8),
Ctrl Trees : 0/0/0, Ctrl ID: 0 (0x0), IR Ctrl ID: 0 (0x0), Ctrl HLI: 0x00000
   P2MP Def MDT ID: 0 (0x0), added: 0, HLI: 0x00000, Cfg: 0/0
   Bidir HLI: 0x00000
```

show mpls mldp database 0x00007

```
Thu Oct  8 15:08:47.137 UTC
 mLDP database
 LSM-ID: 0x00007  Type: MP2MP  Uptime: 01:22:20
   FEC Root           : 100.1.255.254 (we are the root)
   Opaque decoded     : [global-id 3]
   Features           : MBB
   Upstream neighbor(s) :
    None
   Downstream  client(s):
    LDP 100.3.255.254:0 Uptime: 00:02:33
       Next Hop         : 31.7.0.1
       Interface        : TenGigE0/1/0/4/2
       Remote label (D) : 24212             Local label (U) : 24236
      PIM MDT            Uptime: 01:22:20
       Egress intf     : Lmdtvrf/104
      Table ID        : IPv4: 0xe0000006 IPv6: 0xe0800006
      HLI             : 0x00007
      Ingress         : Yes
      Local Label     : 24002 (internal)
```

Equal Cost Multipath Redirect

Equal Cost Multipath (ECMP) redirect provides an automated mechanism to direct multicast traffic onto links with available bandwidth, while avoiding heavily-loaded links in the ECMP, in a multi-ring backbone network. ECMP redirect efficiently utilizes available bandwidth in a backbone network composed of ECMP paths for multicast. ECMP redirect enables routers to automatically allocate multicast traffic to members of an ECMP path, taking into consideration the bandwidth utilization of each member link. ECMP redirect works on native IP multicast infrastructure with PIM.

The network contains multiple optical rings in the backbone forming, equal cost multiple paths (ECMPs), where TV multicast traffic is dominant. With the PIM implementation, a router uses the reverse-path forwarding (RPF) procedure to select an upstream interface and a PIM neighbor to build the forwarding state. This does not allow a desired spread of traffic among the ECMP member links; also it does not avoid either under-utilization of some of the member links, or traffic overflows on other member links. It is not possible to rely on manual allocation because traffic flows cannot scale when more traffic is added and more bandwidth is provisioned through the addition of member links in the ECMPs. With this set as the background, ECMP provides an automated mechanism to direct multicast traffic onto links with available bandwidth and avoids heavily-loaded links in ECMP.

In short, ECMP redirect:

- Steers multicast traffic by policy-based RPF path selection at the downstream nodes.
- Steers multicast traffic at the upstream nodes by triggering PIM ECMP Redirect messages to downstream nodes.

With the ECMP Redirect mechanism, upstream routers use a new PIM ECMP Redirect message to instruct downstream routers on how to tie-break among the upstream neighbors. The PIM ECMP Redirect message conveys the tie-break information, based on metrics selected.

The Bandwidth Aware RPF Concept

The PIM bandwidth aware RPF achieves the following with ECMP links:

- Selects the RPF link based on the bandwidth.
- Avoids putting traffic for the same (S,G) onto multiple ECMP links resulting in a duplicate use of bandwidth.
- Allocates fixed bandwidth amount for multicast traffic on each link.

  For example: A scenario where 8 Gbps out of 10 Gbps link for multicast is configured as threshold. When the configured threshold is exceeded, an alarm is issued for the condition, and continues to accept and forward the multicast traffic. When the maximum bandwidth is exceeded, another alarm is issued and drops excessive joins to protect the existing traffic flows from traffic drops.

Enhancements in PIM Protocol Process

The following enhancements are made in the PIM protocol process:

- A route-policy is configured to map TV channels to the used predetermined bandwidth.
- PIM is configured for the set of interfaces forming the PIM ECMP bundle.
- All PIM routers connected to ECMP bundle snoops all PIM join or prune traffic to inventory the link usage information. The bandwidth usage information is used to guide RPF selection for new multicast streams.
- A PIM router sending a join to a member link of the PIM ECMP bundle selects based on the highest available bandwidth to pick a member link and RPF interface or neighbor.
- Upstream PIM routers notify downstream routers by ECMP redirect messages so that all downstream nodes use same transit network or media to join the upstream router.

PIM rpf-redirect Bundle

PIM rpf-redirect bundle is used to logically group PIM interfaces under one logical container. All bandwidth-aware RPF policies and ECMP Redirect processing is done within the scope of rpf-redirect bundle. Many of such bundles can be created for PIM and, at a point of time, an interface can be part of one bundle.

PIM ECMP Redirect Message Processing

These two changes are introduced in PIM messages:

- **PIM ECMP Redirect Hello option**: PIM ECMP Redirect Hello option is generated when interfaces are configured as part of rpf-redirect bundles.
- **PIM ECMP Redirect message**: PIM ECMP redirect logic activates only if all PIM neighbors on the ECMP bundle links are ECMP Redirect capable. ECMP Redirect message is triggered on receiving PIM join from non-desired outgoing interface. The upstream nodes on member links of PIM rpf-redirect bundle generate ECMP Redirect messages. This message is generated to steer joins of (*, G) or (S, G) on interfaces where these channels are already present. This mechanism helps in allocating multicast traffic evenly on transit networks between such upstream and downstream nodes.

PIM RPF Redirect Mechanism

The following is the sequence of actions involved in a PIM RPF redirect mechanism:

1. When upstream nodes see a S-G join and if there is an existing a same SG join on other interface in same rpf-redirect bundle, the upstream nodes generate an ECMP redirect message in response to processing this join. These ECMP redirect messages are generated periodically on processing joins. These periodic messages are useful to recreate RPF paths when downstream nodes go though an HA event.
2. Upstream nodes generate an ECMP redirect message after verifying that bandwidth is available for the channel on the interface where it is trying to move the channel. If no bandwidth is available, an ECMP redirect message is not generated. However, a syslog is generated to notify this link-overload situation.
3. When a Join inventory changes on an upstream node, the joins are not re-balanced. This ensures simplicity and avoids unnecessary RPF changes on the downstream nodes.
4. ECMP Redirect message generation is preserved in a checkpoint store to handle HA scenarios.

Snooping of PIM Joins on ECMP Links

Due to the population of different group memberships downstream of each PIM routers on the ECMP bundle links, the bandwidth usage on an ECMP bundle link can be contributed by traffic serving other routers. Each of the router performing the ECMP redirect logic snoops all the PIM joins on the ECMP bundle links, including those not directed to that router. As a result, when PIM rpf-redirect is enabled on a router, an (S, G) join is created for the PIM join initiated by other routers, and targeted at other routers, even if the local router has no downstream receivers for such an (S, G) join. These cache entries contain a NULL oif lists, and serve as an aid to inventory the bandwidth usage on the ECMP member links. These entries are refreshed at the PIM join intervals are expired, if not refreshed. The bandwidth usage is updated when such (S, G) is created or deleted.

In addition, ECMP redirect messages are cached on downstream nodes. These cached entries are treated with high priority (over policies) to steer multicast traffic RPF paths. If the bandwidth is not available on the RPF paths as directed by ECMP-redirect message, the Join is ignored. This can be a transient or link over-utilization situation and error syslogs are generated. In a stable state, both the upstream and downstream routers have a consistent view of bandwidth usage and ECMP-redirect messages are not generated.

TV Traffic Profile through Route Policy

Encoding TV traffic yields packets streams with varying bandwidth, the average of these streams is monitored and managed at the encoding side. This helps in administratively configuring the TV channel address to bandwidth usage mapping inside a route policy, such that a router needs to look up this database to determine how much bandwidth a TV channel consumes. There is small number of different types of TV traffic streams, such as high definition, standard definition; therefore, policy configuration is not a tedious task. Also, such policy maps loaded in routers are simple mechanism to monitor bandwidth usage of links as compared to actively polling traffic stats from platform.

Example: The route policy configured for TV channels in address ranges (10,192.X.X, 232.1.Y.Y) with 6 Mbps rate, and (10.192.X.X, 232.2.Y.Y) with 16 Mbps rates: SD channels take 6 Mbps, and HD channels take 16 Mbps

```
route-policy EAST_COAST_CHANNELS
  if destination in (232.1.0.0/16 le 32) and source in (10.192.0.0/16 le 32) then
    set weight 6000   ! Kbps
    endif
       end-policy

 route-policy WEST_COAST_CHANNELS
   if destination in (232.2.0.0/16 le 32) and source in (10.192.0.0/16 le 32) then

    set weight 16000   ! Kbps
     endif
      end-policy
```

A route policy like the one in the above example have only 'S and G' statements and *, G match statement will be an invalid configuration. In other words, only SGs will contribute to bandwidth calculations and *,Gs are ignored.

The RPF path is decided in the following manner:

1. The rpf-redirect bundle maintains a sorted list of its member interfaces in the order of least-used to heavily-used bandwidth. When an S,G join matches a policy under this rpf-redirect bundle, the first interface in the sorted is selected as the RPF path. After processing this join, the interface list is sorted again and the least-used interface becomes the first node.
2. When none of the RPF paths has no bandwidth available, drops the join and generates a syslog message.
3. When an ECMP redirect message is processed, and if the bandwidth is not available on a new RPF path, joins are ignored and a warning syslog is generated.
4. SG RPF paths are preserved in checkpoint store to handle HA scenarios.

Configuring Route Policy for Static RPF

SUMMARY STEPS

1. **configure**
2. **router static**
3. **address-family**[**ipv4** |][**multicast** | **unicast**]*destination prefix interface-typeinterface-path-id*
4. **exit**
5. **route-policy***policy-name*
6. **set rpf-topology** *policy-name***address-family**[**ipv4** |]**multicast** | **unicasttopology***name*
7. **end route-policy**
8. **router pim address-family**[**ipv4** |]
9. **rpf topology route-policy***policy-name***pim policy**

DETAILED STEPS

	Command or Action	Purpose
Step 1	**configure**	
Step 2	**router static** **Example:** `RP/0/RP0/CPU0:router(config) # router static`	Enables a static routing process.
Step 3	**address-family**[**ipv4** |][**multicast** | **unicast**]*destination prefix interface-typeinterface-path-id* **Example:** `RP/0/RP0/CPU0:router(config-static) # address-family ipv4 multicast 202.93.100.4/ 32 202.95.1.1`	Configures the ipv4 multicast address-family topology with a destination prefix.
Step 4	**exit** **Example:** `RP/0/RP0/CPU0:router(config-ipv4-afi) # exit`	Exits from the address family configuration mode.
Step 5	**route-policy***policy-name* **Example:** `RP/0/RP0/CPU0:router(config) # route-policy r1`	Configures the route policy to select the RPF topology.
Step 6	**set rpf-topology** *policy-name***address-family**[**ipv4** |]**multicast** | **unicasttopology***name* **Example:** `RP/0/RP0/CPU0:router(config-rpl) # set rpf-topology p1 ipv4 multicast topology t1`	Configures the PIM rpf-topology attributes for the selected multicast address-family.

Step 7	**end route-policy**
	Example:
	`RP/0/RP0/CPU0:router(config-rpl) #  end route-policy r1`
		Ends the route policy.

Step 8	**router pim address-family**[**ipv4** |] **Example:** `RP/0/RP0/CPU0:router(config) # router pim address-family ipv4`	Enters the PIM address-family configuration sub-mode.
Step 9	**rpf topology route-policy***policy-name***pim policy** **Example:** `RP/0/RP0/CPU0:router(config) # rpf topology route-policy r1 pim policy`	Selects the RPF topology for the configured route-policy.

Native Multicast Collapsed Forwarding

The collapsed forwarding (COFO) feature supports native multicast traffic.

COFO multicast routing supports:

- IPv4 and IPv6
- SM and SSM
- Static RP, BSR and Auto-RP

PIM Bidirectional with COFO native multicast is not supported.

Multicast Traffic Replication Scenerio

The multicast traffic replication depends on the incoming and outgoing interface. The supported replication scenerio is listed below:

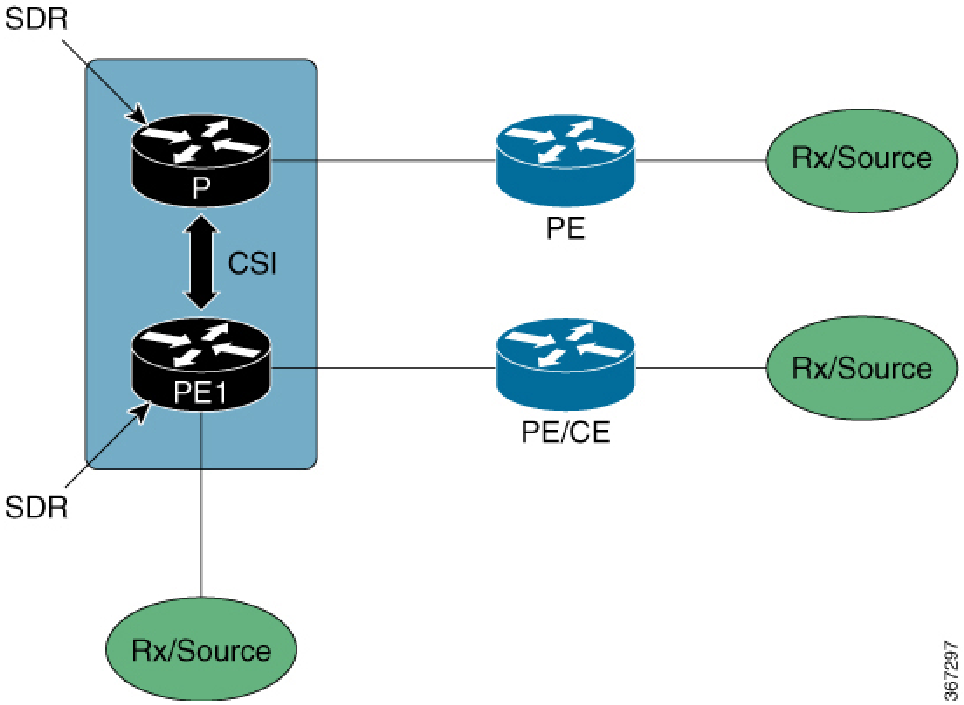
- Ingress traffic from physical interface and outgoing interface is either CSI or physical.
- Ingress traffic from CSI interface and outgoing interface is physical.

The non supported replication scenerio is listed below:

- CSI to CSI interface
- To multiple CSI on same MSE (Multi Service Edge)
- To CSI and non-CSI (to another SDR) on same MSE

Configure Native Multicast with COFO

Consider below topology where two SDRs (P and PE1) are connected to PE and PE/CE routers. The SDRs are connected through CSI interface (CSI1).



Enable multicast routing on CSI interface (CSI1):

```
multicast-routing
 address-family ipv4
  interface CSI1
  accounting per-prefix
 !
 address-family ipv6
  interface CSI1
  accounting per-prefix
```

Verification

Check if CSI interface is in Inlist (incoming interface) on the P router:

```
RP/0/RP0/CPU0:P#sh mrib route 226.1.1.1 detail
(*,226.1.1.1) Ver: 0x56376418 RPF nbr: 192.1.1.1 Flags: C RPF, FGID: 27120, -1, -1 ,
Retry_flags:0
  Up: 01:07:22
  Incoming Interface List
    CSI1 Flags: A, Up: 01:07:22
  Outgoing Interface List
    Bundle-Ether5 (0/3/1) Flags: F NS, Up: 01:07:22
((2.8.1.2,226.1.1.1) Ver: 0x5ffc071a RPF nbr: 192.1.1.1 Flags: RPF, FGID: 150865, -1, -1 ,
 Retry_flags:0
  Up: 01:05:59
  Incoming Interface List
    CSI1 Flags: A, Up: 01:05:59
  Outgoing Interface List
    Bundle-Ether5 (0/3/1) Flags: F NS, Up: 01:05:59
```

Check if CSI interface is in Olist (outgoing interface) on the PE router. Also note down the FGID value for CSI1 interface:

```
RP/0/RP0/CPU0:PE#sh mrib route 226.1.1.1 detail
(*,226.1.1.1) Ver: 0x6db746fc RPF nbr: 2.2.2.2 Flags: C RPF, FGID: 204324, -1, -1 ,
Retry_flags:0
  Up: 01:00:42
  Incoming Interface List
    Decapstunnel0 Flags: A NS, Up: 01:00:42
  Outgoing Interface List
    CSI1 Flags: F NS, Up: 01:00:42
    Bundle-Ether7 (0/1/1) Flags: F NS, Up: 00:58:29
    TenGigE0/0/0/0/0 Flags: F NS LI, Up: 01:00:30
    HundredGigE0/6/0/7 Flags: F NS LI, Up: 00:57:21
(2.8.1.2,226.1.1.1) Ver: 0x6cd0bc19 RPF nbr: 2.8.1.2 Flags: L RPF, FGID: 53285, -1, -1 ,
Retry_flags:0
  Up: 01:00:10
  Incoming Interface List
    TenGigE0/1/0/1/8 Flags: A, Up: 01:00:10
  Outgoing Interface List
    CSI1 Flags: F NS, Up: 01:00:10
    Bundle-Ether7 (0/1/1) Flags: F NS, Up: 00:58:29
    TenGigE0/0/0/0/0 Flags: F NS, Up: 01:00:10
    HundredGigE0/6/0/7 Flags: F NS, Up: 00:57:19
```

Check the FGID 53285 info to view the member nodeset:

```
RP/0/RP0/CPU0:PE#sh mrib fgid info 53285
FGID information
FGID (type)     : 53285 (Primary)
Context         : IP (0xe0000000, 2.8.1.2, 226.1.1.1/32)
Members[ref]    :  0/0/0[1] 0/1/1[1] 0/3/1[1] 0/6/1[1]
FGID bitmap   :0x0000002000080081  0x0000000000000000  0x0000000000000000  0x0000000000000000
0x0000000000000000  0x0000000000000000  0x0000000000000000  0x0000000000000000
0x0000000000000000  0x0000000000000000  0x0000000000000000  0x0000000000000000
0x0000000000000000  0x0000000000000000  0x0000000000000000  0x0000000000000000FGID chkpt
context valid : TRUE
FGID chkpt context :
                table_id 0xe0000000 group 0xe2010101/32 source 0x02080102
FGID chkpt info : 0x23000000

Fgid in batch         : NO
Secondary node count  : 0


In PI retry list :NO
```

You can see the collapsed member nodeset in the below command output:

```
RP/0/RP0/CPU0:PE#sh mrib cofo ip-multicast 226.1.1.1/32
MRIB Collapsed Forwarding DB -- IP Multicast Info:
(*,226.1.1.1)
    Origin: REMOTE
    Receive Count: 1
    Last Received: Thu Oct 12 03:18:55 2017
    Nodeset: 0/3/1
(2.8.1.2,226.1.1.1)
    Origin: REMOTE
    Receive Count: 1
    Last Received: Thu Oct 12 04:13:30 2017
    Nodeset: 0/3/1
```

Check the MRIB COFO database to verify local and remote multicast routes:

```
RP/0/RP0/CPU0:PE#sh mrib cofo summary
MRIB COFO DB Summary:
```

```
Type                                  |   Local   |   Remote
--Number of (*,G) entries |          0|            100|
Number of (S,G) entries |          0|                 0|
Number of label entries |             0|              0|
Number of encap entries |           0|            0|
```

In the above output, the *Local* entries show local routes and *Remote* entries show collapsed route from remote SDR.

Collapsed Forwarding for P2MP-TE GTM Profile

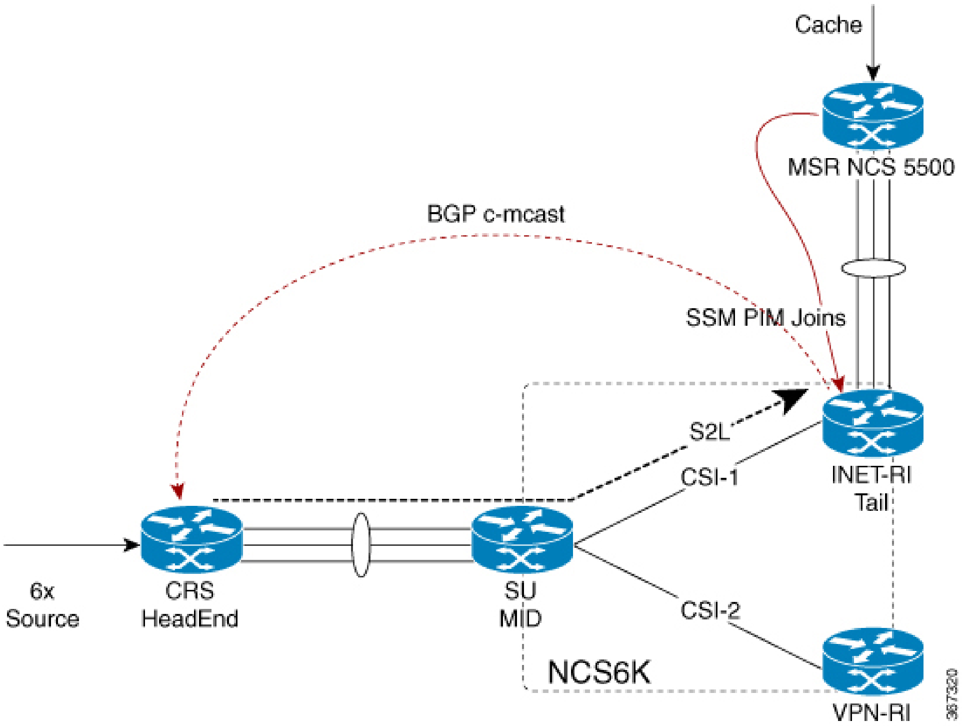
In collapsed forwarding (COFO), inter SDR traffic is handled by the internal fabric itself without requiring the external cables. A newly created SDR interface functions as a point-to-point virtual interface, connecting SDR routers in the system. This virtual interface that connects two SDRs to each other is known as cross SDR interconnect (CSI) interface.

The collapsed forwarding feature supports Multicast over Point-to-Multipoint-Traffic Engineering (P2MP-TE) GTM profile.

Note

- In Cisco IOS XR Release 6.4.1, COFO P2MP TE GTM supports only egress (tail node) functionality.
- When a protected link (physical/bundle) is shut down in SU, TE FRR kicks in at SU and VPN traffic is protected. When no shut of a protected link is performed, IGP comes up first and VPN switches to NNH as soon as IGP restores on the protected link. However, TE FRR takes more time to reoptimize and until TE re-optimization completes, traffic over TE tunnel drops. TE tunnel starting from VPN and FRR is protected on SU.

The figure shows collapsed forwarding topology for P2MP-TE GTM.



An auto tunnel is formed from headend to the tail, a source-to-leaf path. The tail will share route details using BGP with headend.

From the tailnode (INET-RI) the label and FGID values are collapsed to the midnode (SU). The traffic from the midnode will make use of these values to reach the correct egress line card of the tailnode.

The other characterstics include:

- P2MP-TE COFO is next-next-hop (NNH) based.
- Collapse of label and FGID information is from tail to mid.

Collapsed Forwarding Global Table Multicast and Egress PE

In the case where global table multicast uses PIM-SM in ASM (Any Source Multicast) mode, the inter-area P2MP service LSP can be used to carry traffic either on a shared (*,G) or a source (S,G) tree.

An egress PE learns the SSM of a multicast stream as a result of receiving IGMP or PIM messages on one of its IP multicast interfaces. This SSM forms the P2MP FEC of the inter-area P2MP service LSP. For each of such a P2MP FEC, a distinct inter-area P2MP service LSP may exist, or multiple FECs may be carried over a single P2MP service LSP using a wildcard (*,*) S-PMSI.

The following are not supported for CSI:

- (*, G)
- Bud node and branch scenario
- P2MP TE GTM profile for IPv6
- P2MP TE headend

Functional Overview of P2MP TE for CSI

Tunnel Head

At the Headend, RSVP-TE signals and establishes a P2MP TE LSP. In order to forward the multicast traffic over the P2MP TE LSP, the multipoint tunnel interface is enabled for multicast forwarding.

Additionally, MPLS and RSVP-TE are activated on each MPLS core-facing interface.

Midpoint

At mid node, no multicast routing protocol runs. Therefore, there is no configuration related to the multicast routing. However, a global configuration is required for activating LMRIB function in the multicast package. This configuration is set at all nodes that process P2MP LSPs.

MPLS and RSVP-TE are activated on each MPLS core-facing interface.

The mid-node signals and establishes a P2MP TE LSP according to the signaling message from the upstream node. This node conducts the label packet replication for forwarding the packet to each downstream node.

Tunnel Tail

At the Tailend, RSVP-TE signals and establishes a P2MP TE LSP. MPLS and RSVP-TE are thus configured on each MPLS core-facing interface.

PIM conduct a special RPF check for the packet that is received from MPLS core.

Instead of the normal IP RPF check, the tail end node checks whether the packet comes from the correct headend, or not. If not, this node drops the received packet.

Configuring Collapsed Forwarding for P2MP-TE GTM Profile

Configuring collapsed forwarding for P2MP-TE GTM involves multicast configuration for:

- Headend
- Midnode
- Tailend

Multicast Configuration on Headend

```
mpls traffic-eng
auto-tunnel p2mp
  tunnel-id min 2001 max 3001
 !
!
multicast-routing
 address-family ipv4
  mdt source Loopback0
  export-rt 701:0
  import-rt 701:0
  rate-per-route
  interface all enable
  accounting per-prefix
  bgp auto-discovery p2mp-te
  !
  mdt default p2mp-te
  mdt data p2mp-te 1000 immediate-switch
 !
```

```
!
router bgp 4134
 nsr
 mvpn
 bgp router-id 1.1.1.1
 bgp log neighbor changes detail
 address-family ipv4 unicast
  redistribute connected
  allocate-label all
 !
 address-family ipv4 multicast
  redistribute connected
 !
address-family ipv4 mvpn
  global-table-multicast
 !
```

Multicast Configuration on Midnode

```
multicast-routing
 address-family ipv4
  interface CSI12
   disable
  !
rate-per-route
accounting per-prefix
 !
```

Multicast Configuration on Tailend

```
multicast-routing
 address-family ipv4
  interface CSI12
   disable
  !
  mdt source Loopback0
  export-rt 701:0
  import-rt 701:0
  maximum disable
  rate-per-route
  interface all enable
  accounting per-prefix
  bgp auto-discovery p2mp-te
   receiver-site
  !
  mdt default p2mp-te
 !
!
router pim
 address-family ipv4
  rpf topology route-policy GTM
  mdt c-multicast-routing bgp
  !
 !
!
route-policy GTM
  set core-tree p2mp-te-default
end-policy
!
router bgp 4134
address-family ipv4 multicast
  redistribute connected
 !
address-family ipv4 mvpn
  global-table-multicast
 !
```

Verification

From the tailnode (INET-RI) the label and FGID values are collapsed to the midnode (SU). The traffic from the midnode will make use of these values to reach the correct egress line card of the tailnode.

To view and verify the collapsed label and FGID info:

1. In the headend-using the tunnel number, find out incoming and outgoing label associated with that tunnel
2. In the midnode-using the tunnel number, find out incoming and outgoing label associated with that tunnel. Use the incoming label to find out FGID

Headend

Use the **show mrib route** command to find out the tunnel that multicast traffic is using. In our topology, the traffic is using tunnel 2516.

Use the **show mrib mpls for tunnels 2516** to find out incoming and outgoing label associated with tunnel 2516:

```
SDR-1#sh mrib mpls for tunnels 2516
Thu Aug  3 13:56:58.415 IST

LSP information (RSVP-TE) :
  Name: ------, Role: Head
    Tunnel-ID: 2516, P2MP-ID: 2516, LSP-ID: 10002
    Source Address: 1.1.1.1, Extended-ID: 1.1.1.1
    Incoming Label       : 24832
    Transported Protocol : IPv4
    Explicit Null        : IPv6 Explicit Null
    IP lookup            : disabled

    Outsegment Info #1 [M/Swap]:
      OutLabel: 24809, NH: 101.12.1.2, IF: BE12001
```

The incoming label is 24832 and outgoing label 24809.

Midnode

Use the **show mrib mpls for tunnels 2516** to find out incoming and outgoing label associated with the tunnel 2516 in the midnode:

```
#sh mrib mpls  for  tunnels 2516
Thu Aug  3 13:56:58.415 IST

LSP information (RSVP-TE) :
  Name: ------, Role: Mid
    Tunnel-ID: 2516, P2MP-ID: 2516, LSP-ID: 10002
    Source Address: 1.1.1.1, Extended-ID: 1.1.1.1
    Incoming Label       : 24809
    Transported Protocol : IPv4
    Explicit Null        : IPv6 Explicit Null
    IP lookup            : disabled
    Outsegment Info #1 [M/Swap]:
      OutLabel: 24833, NH: 192.168.12.2, IF: CSI12
```

The incoming label is 24809 and outgoing label is 24833. The outgoing label of headend is now the incoming label for midnode.

Use the **sh mpls forwarding labels 24809 hardware ingress detail location** command to find out FGID:

```
#sh mpls  forwarding labels 24809 hardware ingress  detail location 0/1/CPU1
Thu Aug  3 13:57:15.023 IST
Local  Outgoing    Prefix             Outgoing     Next Hop        Bytes
Label  Label       or ID              Interface                    Switched
------ ----------- ------------------ ------------ --------------- ------------
24935              P2MP TE: 2516

       24833       P2MP TE: 2516      CS12         192.168.12.2    N/A

HW Walk:
NPU #1
  LEAF: 0x241859e0 (Offset: 0)
  HW: 0x000006 00000000 00000000 10f25340
  entrytype  : P2MP         leaf       : 0            dccheck    : 0
  islabelptr : 0            islabel    : 0            bgppa      : 0
  label/array: 0            flowtag    : 0
  baoId      : 0            prefixlen  : 0            qosgroup   : 0
  nextptr    : 0x10f2534    numentries : 1

      MPLS P2MP RX: 0x10f2534 (Offset: 0)
      HW: 0x6167092 00000100 04dcd000 04dcc000
      label      : 24935        label valid: 1            stats_ptr  : 0
      fgid       : 0x4dcc       bkup fgid  : 0x4dcd
      tunnel uidb: 0            cofo enable: 1            peek ctrl  : 0
      punt oam   : 0            punt data  : 0            drop       : 0
      tluid      : MPLS P2MP RX

      MPLS P2MP RX EXT: 0x10f2535 (Offset: 0)
      HW: 00000000 00000000 00000000 00000000
      frrslotmskh: 0            frrslotmskl: 0
      frrslicemsk: 0            frrslicemsk: 0

      MPLS P2MP GLOBAL FRR MASK: 0x1083dd8 (Offset: 0)
      HW: 00000000 00000000 00000000 00000000
      frrslotmskh: 0            frrslotmskl: 0
      frrslicemsk: 0            frrslicemsk: 0

      COFO: RX MPLS P2MP OLIST HEAD: 0x10f2536 (Offset: 0)
      HW: 0x000030 10f217a0 00000000 00000000
      olist_ptr  : 0x10f217a    olist_empty: 0            l3fib_tlu_i: MPLS OLIST HEAD
      v4_exp_null: 0            v6_exp_null: 1

          COFO:RX MPLS P2MP OLIST ENTRY: 0x10f217a (Offset: 0)
          HW: 0x610102c 0000000c 00000008 10f2e500
          olist_ptr  : 0            olist_empty: 1            l3fib_tlu_i: COFO MPLS OLIST
 ENTRY
          label      : 24833        label_valid: 1            label_stats: 0
          label_stats: 0            csi index  : 12           next_ptr   : 0x10f2e50

              COFO NFGID DB: 0x10f2e50 (Offset: 0)
              HW: 0x000017 00000000 04600000 74240000
              l3fib_tlu_i: COFO NFGID DB   primary_fgi: 117780       backup_fgid: 475712

              MPLS P2MP RX EXT: 0x10f2e51 (Offset: 0)
              HW: 00000000 00000000 00000000 00000000
              frrslotmskh: 0            frrslotmskl: 0
```

```
frrslicemsk: 0            frrslicemsk: 0

MPLS P2MP GLOBAL FRR MASK: 0x10ee928 (Offset: 0)
HW: 00000000 00000000 00000000 00000000
frrslotmskh: 0            frrslotmskl: 0
frrslicemsk: 0            frrslicemsk: 0
```

MVPN GRE over PWHE with CSI

MVPN GRE over PWHE is supported on CSI interface.

The Multicast VPN (MVPN) feature provides the ability to support multicast over a Layer 3 VPN. Whereas, Pseudowire Headend (PWHE) allows termination of access pseudowires (PWs) into a Layer 3 (VRF or global) domain or into a Layer 2 domain.

Note From Cisco IOS XR, Release 7.1.1 both IPv4 and IPv6 are supported on PE-CE multicast over PWHE interfaces.

Restrictions

- Only SSM is supported on PE-CE multicast
- Only IPv4 SM is supported on provider multicast
- Does not support SM on PE-CE multicast
- Does not support ISSU

Configuration Example

```
interface PW-Ether1 vrf vrf1
 ipv4 address 192.0.2.1 255.255.255.252
 ipv6 address 2001:DB8:1::1/32
 attach generic-interface-list Bundle311
!
```

Configuration Examples for Implementing Multicast Routing on Software

This section provides the following configuration examples:

DNS-based SSM Mapping: Example

The following example illustrates DNS-based SSM Mapping configuration.

```
multicast-routing
 address-family ipv4
```

```
  nsf
  mdt source Loopback5
  maximum disable
  interface all enable
  accounting per-prefix
 !
 address-family ipv6
  nsf
  maximum disable
  interface all enable
  accounting per-prefix
 !
 vrf p11_1
  address-family ipv4
   ssm range ssm_acl
   interface all enable
   mdt default ipv4 235.1.1.1
  !



ipv4 access-list ssm_acl
 10 permit ipv4 225.11.1.0 0.0.0.255 any
 20 permit ipv4 225.11.2.0 0.0.0.255 any
!

router mld
 vrf p11_1
  ssm map query dns


router igmp

 !
 vrf p11_1
  ssm map query dns
 !


domain vrf p11_1 name-server 100.1.1.2
domain multicast cisco.com
domain name-server 10.10.10.1
```

Preventing Auto-RP Messages from Being Forwarded on Software: Example

This example shows that Auto-RP messages are prevented from being sent out of the interface 0/3/0/0. It also shows that access list 111 is used by the Auto-RP candidate and access list 222 is used by the **boundary** command to contain traffic on interface 0/3/0/0.

```
ipv4 access-list 111
 10 permit 224.1.0.0 0.0.255.255 any
 20 permit 224.2.0.0 0.0.255.255 any
!
!Access list 111 is used by the Auto-RP candidate.
!
ipv4 access-list 222
 10 deny any host 224.0.1.39
 20 deny any host 224.0.1.40
!
!Access list 222 is used by the boundary command to contain traffic (on /3/0/0) that is
sent to groups 224.0.1.39 and 224.0.1.40.
```

```
!
router pim
 auto-rp mapping-agent loopback 2 scope 32 interval 30
 auto-rp candidate-rp loopback 2 scope 15 group-list 111 interval 30
multicast-routing
 interface /3/0/0
 boundary 222
!
```

Inheritance in MSDP on Software: Example

The following MSDP commands can be inherited by all MSDP peers when configured under router MSDP configuration mode. In addition, commands can be configured under the peer configuration mode for specific peers to override the inheritance feature.

- **connect-source**
- **sa-filter**
- **ttl-threshold**

If a command is configured in both the router msdp and peer configuration modes, the peer configuration takes precedence.

In the following example, MSDP on Router A filters Source-Active (SA) announcements on all peer groups in the address range 226/8 (except IP address 172.16.0.2); and filters SAs sourced by the originator RP 172.16.0.3 to 172.16.0.2.

MSDP peers (172.16.0.1, 172.16.0.2, and 172.17.0.1) use the loopback 0 address of Router A to set up peering. However, peer 192.168.12.2 uses the IPv4 address configured on the interface to peer with Router A.

Router A

```
!
ipv4 access-list 111
 10 deny ip host 172.16.0.3 any
 20 permit any any
!

ipv4 access-list 112
 10 deny any 226.0.0.0 0.255.255.255
 30 permit any any
!
router msdp
 connect-source loopback 0
 sa-filter in rp-list 111
 sa-filter out rp-list 111
 peer 172.16.0.1
!
peer 172.16.0.2
 sa-filter out list 112
!
peer 172.17.0.1
!
peer 192.168.12.2
 connect-source /2/0/0
!
```

Configuring Route Policy for Static RPF: Example

```
router static
 address-family ipv4 multicast
  202.93.192.74 /32 202.40.148.11

!
route-policy pim-policy
 set rpf-topology ipv4 multicast topology default

end-policy
!
router pim
 address-family ipv4
  rpf topology route-policy pim-policy
```

Additional References

Related Documents

Related Topic	Document Title
Multicast command reference document	*Multicast Command Reference for Cisco NCS 6000 Series Routers*
Getting started material	
Modular quality of service command reference document	
Routing command reference and configuration documents	*Routing Command Reference for Cisco NCS 6000 Series Routers* *Routing Configuration Guide for Cisco NCS 6000 Series Routers*
Information about user groups and task IDs	*System Security Configuration Guide for Cisco NCS 6000 Series Routers*