Nokia Service Router Linux
7730 Service Interconnect Router
7220 Interconnect Router
7250 Interconnect Router

Release 26.3

Multicast Guide

3HE 22253 AAAA TQZZA
Edition: 01
March 2026

Nokia is committed to diversity and inclusion. We are continuously reviewing our customer documentation and consulting with standards bodies to ensure that terminology is inclusive and aligned with the industry. Our future customer documentation will be updated accordingly.

This document includes Nokia proprietary and confidential information, which may not be distributed or disclosed to any third parties without the prior written consent of Nokia.

This document is intended for use by Nokia's customers ("You"/"Your") in connection with a product purchased or licensed from any company within Nokia Group of Companies. Use this document as agreed. You agree to notify Nokia of any errors you may find in this document; however, should you elect to use this document for any purpose(s) for which it is not intended, You understand and warrant that any determinations You may make or actions You may take will be based upon Your independent judgment and analysis of the content of this document.

Nokia reserves the right to make changes to this document without notice. At all times, the controlling version is the one available on Nokia's site.

No part of this document may be modified.

NO WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO ANY WARRANTY OF AVAILABILITY, ACCURACY, RELIABILITY, TITLE, NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE, IS MADE IN RELATION TO THE CONTENT OF THIS DOCUMENT. IN NO EVENT WILL NOKIA BE LIABLE FOR ANY DAMAGES, INCLUDING BUT NOT LIMITED TO SPECIAL, DIRECT, INDIRECT, INCIDENTAL OR CONSEQUENTIAL OR ANY LOSSES, SUCH AS BUT NOT LIMITED TO LOSS OF PROFIT, REVENUE, BUSINESS INTERRUPTION, BUSINESS OPPORTUNITY OR DATA THAT MAY ARISE FROM THE USE OF THIS DOCUMENT OR THE INFORMATION IN IT, EVEN IN THE CASE OF ERRORS IN OR OMISSIONS FROM THIS DOCUMENT OR ITS CONTENT.

Table of contents

1 About this guide

This guide provides an overview of multicast routing concepts and configuration examples for Internet Group Management Protocol (IGMP), Multicast Listener Discovery (MLD), Protocol Independent Multicast (PIM), and Multicast Source Discovery Protocol (MSDP) on Nokia Service Router Linux (SR Linux).

The multicast functionality is supported on 7250 IXR Gen 2/2c+, 7730 SXR, and 7220 IXR-D*x* platforms. For platform support details, see Supported platforms.

Note:
This manual covers the current release and may also contain some content that will be released in later maintenance loads. See the *SR Linux Release Notes* for information about features supported in each load.

Configuration and command outputs shown in this guide are examples only; actual displays may differ depending on supported functionality and user configuration.

1.1 Precautionary and information messages

The following are information symbols used in the documentation.

DANGER: Danger warns that the described activity or situation may result in serious personal injury or death. An electric shock hazard could exist. Before you begin work on this equipment, be aware of hazards involving electrical circuitry, be familiar with networking environments, and implement accident prevention procedures.

WARNING: Warning indicates that the described activity or situation may, or will, cause equipment damage, serious performance problems, or loss of data.

Caution: Caution indicates that the described activity or situation may reduce your component or system performance.

Note: Note provides additional operational information.

Tip: Tip provides suggestions for use or best practices.

1.2 Conventions

The Nokia SR Linux documentation uses the following command conventions:

- **Bold** type indicates a command that the user must enter.
- Input and output examples are displayed in `Courier` text.
- A vertical bar (|) indicates a mutually exclusive argument.
- Square brackets ([ ]) indicate optional elements.

- Braces ({ }) indicate a required choice. When braces are contained within square brackets, they indicate a required choice within an optional element.
- *Italic* type indicates a variable.

The following table outlines platform grouping conventions used in the SR Linux documentation suite.

Note: Some platforms in the 7250 IXR support mixed systems. For more information about mixed system support, see "Chassis types" in the *Configuration Basics Guide*.

Table 1: Platform grouping legend

Platform group	Description
7215 IXS	7215 IXS-A1
7220 IXR	All 7220 IXR platforms
7220 IXR-D*x*	7220 IXR-D1, 7220 IXR-D2, 7220 IXR-D2L, 7220 IXR-D3, 7220 IXR-D3L, 7220 IXR-D4, 7220 IXR-D5
7220 IXR-H*x*	7220 IXR-H2, 7220 IXR-H3, 7220 IXR-H4, 7220 IXR-H4-32D, 7220 IXR-H5-32D, 7220 IXR-H5-64D, 7220 IXR-H5-64O
7250 IXR[1]	7250 IXR platforms
7250 IXR Gen 2	7250 IXR-6, 7250 IXR-10
7250 IXR Gen 2c+	7250 IXR-6e with IMM2, 7250 IXR-10e with IMM2, 7250 IXR-X1b, 7250 IXR-X3b
7250 IXR Gen 3	7250 IXR-6e with IMM3, 7250 IXR-10e with IMM3, 7250 IXR-18e, 7250 IXR-X4
7250 IXR-6e/10e (mixed system)	7250 IXR-6e (mixed system)[2], 7250 IXR-10e (mixed system)[2]
7730 SXR	7730 SXR-1-32D, 7730 SXR-1d-32D, 7730 SXR-1x-44S

[1] References to the 7250 IXR platform group may be appended with (including mixed systems) or (excluding mixed systems) to indicate mixed system support.

[2] References to this platform as part of 7250 IXR (mixed system) indicate mixed system support of 7250 IXR Gen 2c+ (IMM2) and 7250 IXR Gen 3 (IMM3). That is, the 7250 IXR-6e and 7250 IXR-10e can hold and support both IMM2 and IMM3 at the same time.

2 What's new

This section lists the changes that were made in this release.

Table 2: What's new in Release 26.3.1

Topic	Location
Multicast Source Discovery Protocol (MSDP) support on 7250 IXR Gen 2/2c+ (**only** for default network instances)	Multicast Source Discovery Protocol (MSDP)

3 Multicast overview

IP multicast offers an efficient solution for many-to-many communication. Delivering unicast datagrams is fairly simple. With unicast, IP packets are sent from a single source to a single recipient. The source inserts the address of the target host in the IP header destination field of an IP datagram; intermediate routers (if present) forward the datagram toward the target in accordance with their respective routing tables.

Sometimes, distribution needs individual IP packets be delivered to multiple destinations (like audio or video streaming broadcasts). Multicast is a method of distributing datagrams sourced from one or more hosts to a set of receivers that may be distributed over different (sub) networks. This makes the delivery of multicast datagrams significantly more complex.

Multicast sources can send a single copy of data using a single address for the entire group of recipients. The routers between the source and recipients route the data using the group address route. Multicast packets are delivered to a multicast group. A multicast group specifies a set of recipients who are interested in a particular data stream and is represented by an IP address from a specified range. Data addressed to the IP address is forwarded to the group members. A source host sends data to a multicast group by specifying the multicast group address in the datagram’s destination IP address. A source does not have to register to send data to a group, nor do they need to be a member of the group.

Routers and Layer 3 switches use the Internet Group Management Protocol (IGMP) to manage membership for a multicast session. When a host wants to receive one or more multicast sessions, it sends a join message for each multicast group it wants to join. When a host wants to leave a multicast group, it sends a leave message.

The SR Linux routers primarily use a combination of the following protocols:

- Internet Gateway Multicast Protocol (IGMP)
- Multicast Listener Discovery (MLD)
- Protocol Independent Multicast (Sparse Mode) (PIM-SM)
- Multicast Source Discovery Protocol (MSDP)

Supported platforms

The multicast functionality is supported on 7250 IXR Gen 2/2c+, 7730 SXR, and 7220 IXR-D*x* platforms as shown in the following table:

Table 3: Multicast platform support

Platforms	Network instance support	IGMP versions	PIM-SM (IPv4/IPv6)	SSM-translation	MLD (v1/v2)	MSDP
7250 IXR Gen 2/2c+	Default only	1,2,3	Yes	v1, v2	MLDv1 with SSM-translation and MLDv2	Yes (**only** default instance)
7730 SXR	Default and non-default	1,2,3	Yes	v1, v2	MLDv1 with SSM-translation and MLDv2	Not supported

Platforms	Network instance support	IGMP versions	PIM-SM (IPv4/IPv6)	SSM-translation	MLD (v1/v2)	MSDP
7220 IXR-D*x*	Default and non-default	1,2,3	Yes	v1, v2	MLDv1 with SSM-translation and MLDv2	Not supported

3.1 Multicast models

SM

Sparse Mode Multicast (SM) is the IP multicast service model defined in RFC 7761, *Host Extensions for IP Multicasting*. An IP datagram is transmitted to a host group, a set of zero or more end-hosts, identified by a single IP destination address within the range 224.0.0.0 to 239.255.255.255 for IPv4 and FF00::/8 for IPv6. End-hosts can join and leave the group at any time, and there is no limitation on their location or number. This model supports multicast groups with arbitrarily many senders. Any end-host can transmit to a host group even if it is not a member. Any Source Multicast (ASM) is represented by `(*, G)` notation. The `*` (wildcard) indicates that the multicast group can receive traffic from any source. `G` represents the multicast group address.

SSM

The Source-Specific Model (SSM) is also explained in RFC 7761 and it allow a receiver to select the group it wants to receive and the sources from which it wants to receive the data. This extension is intended to give administrators greater control over senders in the multicast network. The source-specific multicast range (232.0.0.0/8) includes addresses reserved for use with the PIM SSM model.

The source-specific model defines a channel identified by an (S, G) pair, where S is a source address and G is an SSM destination address. In contrast to the ASM model, SSM only provides network-layer support for one-to-many delivery.

The SSM service model attempts to alleviate the following deployment problems that SM has presented:

- **address allocation**

  SSM defines channels on a per-source basis. For example, the channel (S1, G) is distinct from the channel (S2, G), where S1 and S2 are source addresses, and G is an SSM destination address. This averts the problem of global allocation of SSM destination addresses and makes each source independently responsible for resolving address collisions for the various channels it creates.

- **access control**

  SSM provides an efficient solution to the access control problem. When a receiver subscribes to an (S, G) channel, it receives data sent only by the source S. In contrast, any host can transmit to an ASM host group. At the same time, when a sender picks a channel (S, G) to transmit on, it is automatically ensured that no other sender is transmitting on the same channel (except in the case of malicious acts such as address spoofing). This makes it harder to spam an SSM channel than an SM multicast group.

- **handling of well-known sources**

  SSM requires only source-based forwarding trees, eliminating the need for a shared tree infrastructure. In terms of the IGMP, PIM-SM, Multicast Source Discovery Protocol (MSDP), and Multiprotocol BGP (MBGP) protocol suite, this implies that neither the Rendezvous Point (RP)-based shared tree infrastructure of PIM-SM nor the MSDP protocol is required. Thus, the complexity of the multicast routing infrastructure for SSM is low, making it viable for immediate deployment. MBGP is still required to distribute multicast reachability information.

3.2 Multicast protocols

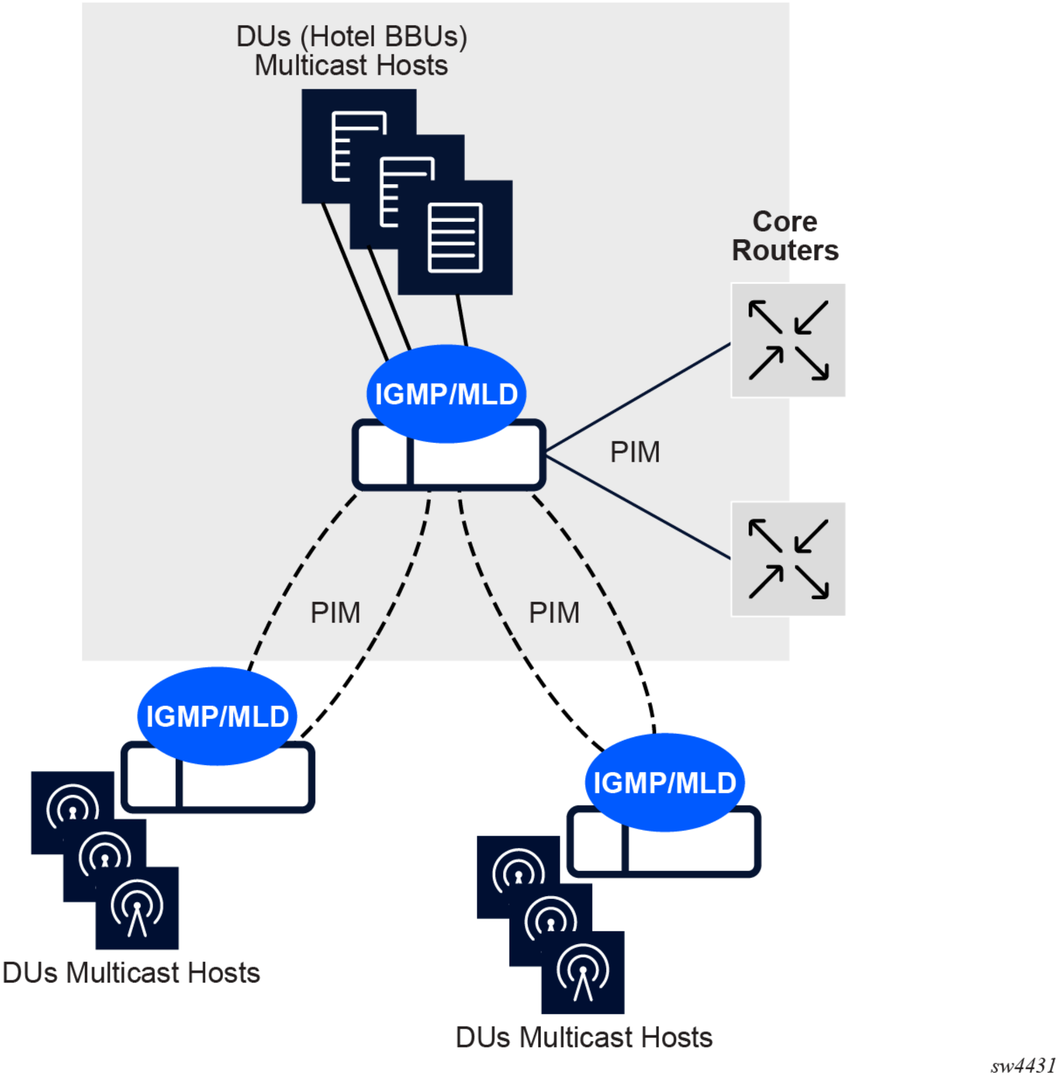
SR Linux supports the following multicast protocols:

Table 4: Multicast protocols

Protocols	Layer	Usage	IPv4/IPv6
IGMP	Layer 3	Handles host and multicast device communication	IPv4
MLD	Layer 3	Handles host and multicast device communication	IPv6
PIM	Layer 3	Manages multicast routing	IPv4/IPv6
MSDP	Layer 2	Provides multiple spanning tree instances to prevent loops in VLAN-enabled networks	NA (operates at Layer 2, independent of IP version)

In the following example, the Base Band Units (BBUs), such as 4G/5G/hotel BBUs, are connected to the cell site router.

Figure 1: Multicast protocols



These BBUs support IGMP or MLD to indicate their intent to join or leave a multicast group, with signaling typically being source-specific multicast (<S, G>).

In some scenarios, older equipment at the cell site, such as cameras, can signal <*, G>. In these scenarios, the <*, G> signal can either be translated to <S, G> via IGMP/MLD translation or PIM ASM with Rendezvous Point (RP) configuration can be used to signal joins and prunes to the RP.

The cell site router translates these IGMP or MLD join or prunes into PIM join or prunes on the Network-to-Network Interface (NNI), signaling the multicast to the source.

3.3 Multicast policies

Multicast traffic can be restricted from specific source addresses by creating routing policies. When a multicast policy is applied to IGMP, MLD, or PIM, the multicast parameters in the policy statement are compared to the (S,G) or (*, G) in the incoming join or prune messages for the protocol. If a match is found, the join or prune message is accepted or denied based on the policy action. The same policy can be applied to multiple protocols.

4 Internet Group Management Protocol (IGMP)

IGMP is the multicast signaling protocol IPv4 hosts and routers use to report their IP multicast group memberships to neighboring multicast routers. A multicast router keeps a list of multicast group memberships for each attached network and a timer for each membership.

Multicast group memberships include at least one member of a multicast group on a specific attached network, not a list of all the members. With respect to each of its attached networks, a multicast router can assume one of two roles, querier or non-querier. There is normally only one querier per physical network.

A querier issues two types of queries: a general query and a group-specific query. General queries are issued to solicit membership information with regard to any multicast group. Group-specific queries are issued when a router receives a leave message from the node it perceives as the last group member remaining on that network segment.

Hosts wanting to receive a multicast session issue a multicast group membership report. These reports must be sent to all multicast -enabled routers.

IGMP support in SR Linux includes the following:

- IGMP versions 1, 2, and 3 are supported.
- Querier support, including timers and relevant parameters.
- Trace options.
- Static membership groups configured under interfaces.
- Support of static <S, G> and <*, G> translation to <S, G>.

4.1 IGMP versions and interoperability requirements

If routers run different versions of IGMP, they negotiate the lowest common version of IGMP that is supported on their subnet and operate in that version.

- **Version 1**

  Specified in RFC 1112, *Host extensions for IP Multicasting*, was the first widely deployed version and the first version to become an Internet standard.

- **Version 2**

  Specified in RFC 2236, *Internet Group Management Protocol, Version 2*, added support for "low leave latency", that is, a reduction in the time required for a multicast router to determine that no members of a specific group are present on an attached network.

- **Version 3**

  Specified in RFC 3376, *Internet Group Management Protocol, Version 3*, adds support for source filtering; that is, the ability for a system to report interest in receiving packets only from specific source addresses, as required to support SSM, or from all but specific source addresses, sent to a particular multicast address.

IGMPv3 must keep state per group per attached network. This group state consists of a filter-mode, a list of sources, and various timers. For each connected network running IGMP, a multicast router records the needed reception state for that network.

4.2 IGMP version transition

SR Linux routers can interoperate with routers and hosts running IGMPv1, IGMPv2, or IGMPv3. *Draft-ietf-magma-igmpv3-and-routing-0x.txt* explores some of the interoperability issues and how they affect the various routing protocols.

IGMP version 3 specifies that if at any point a router receives an older version query message on an interface, it must immediately switch into a compatibility mode with that earlier version. Because none of the previous versions of IGMP are source aware, should this occur and the interface switch to Version 1 or 2 compatibility mode, any previously learned group memberships with specific sources (learned from the IGMPv3 specific INCLUDE or EXCLUDE mechanisms) must be converted to non-source specific group memberships. The routing protocol then treats this as if no EXCLUDE definition is present.

4.3 Source Specific Multicast (SSM) groups

IGMPv3 allows a receiver to join a group while specifying that it only wants to receive traffic from a particular source. If a receiver does this, and no other receiver on the LAN requires traffic from all sources for that group, then the designated router (DR) can omit performing a (*, G) join to set up the shared tree, and instead issue a source-specific (S, G) join only.

The range of multicast addresses from 232.0.0.0 to 232.255.255.255 is set aside for source-specific multicast in IPv4. For groups in this range, receivers should only issue source-specific IGMPv3 joins. If a PIM router receives a non-source-specific join for a group in this range, it should ignore it.

An SR Linux PIM router must silently ignore a received (*, G) PIM join message where G is a multicast group address from the multicast address group range that has been explicitly configured for SSM. This occurrence should generate an event. If configured, the IGMPv2 request can be translated into IGMPv3. The router allows for the conversion of an IGMPv2 (*, G) request into an IGMPv3 (S, G) request based on manual entries. A maximum of 32 SSM ranges is supported.

IGMPv3 also allows a receiver to join a group while specifying that it wants to receive traffic only if it does not originate from specific sources. In this case, the DR performs a (*, G) join as usual but can combine this with a prune for each source the receiver does not want to receive

4.4 Query messages

The IGMP query source address is configurable at two hierarchical levels. It can be configured globally at each router instance IGMP level and can be configured individually at the group-interface level. The group-interface level overrides the source IP address configured at the router instance level.

By default, subscribers with IGMP policies send IGMP queries with an all zero SRC IP address (0.0.0.0). However, some systems only accept and process IGMP query messages with non-zero SRC IP addresses. This feature allows the Broadband Network Gateway (BNG) to inter-operate with those systems.

4.5 IGMP configuration

The routers use IGMP to manage membership for a multicast session. IGMP is configured by enabling `igmp admin-state` under the `network-instance protocols` context.

The `igmp` protocol allows the configuration of IGMP parameters in both the default and non-default network instance types, with the ability to apply specific configurations to individual interfaces.

When enabled, at least one interface must be specified in the IGMP context, as IGMP is an interface function. Traffic can only flow away from the router to an IGMP interface and to and from a PIM interface. A router directly connected to a source must have PIM enabled on the interface to that source. The traffic travels in a network from PIM interface to PIM interface and arrives finally on an IGMP enabled interface.

In a typical multicast operation, when a host wants to receive traffic for a multicast group, it sends an IGMP membership report (acting as a join message) to the multicast group address. This report notifies the multicast router that the host is interested in joining the group and receiving its traffic. When the host no longer wants to receive multicast traffic, it may send an IGMP leave message (in IGMPv2 or later) to signal that it is leaving the group.

A multicast router keeps a list of multicast group memberships for each attached network and an interval timer for each membership. It periodically sends IGMP queries to check whether hosts still wants to remain part of each group. Hosts still wanting to receive traffic must respond with an updated IGMP membership report. The router maintains an interval for each membership and removes memberships if it no longer gets the required reports.

4.5.1 Basic IGMP configuration

Procedure

IGMP is configured by enabling `igmp admin-state` under the `network-instance protocols` context.

Example: Configuring IGMP for a default network-instance

The following example shows a configuration of IGMP for a default network instance (`srl_default_instance`). It includes enabling IGMP, configuring basic IGMP parameters, such as configuring IGMP on an interface, specifying the IGMP version for an interface, and setting up the query interval.

```
--{ +* candidate shared default }--[  ]--
# info with-context network-instance srl_default_instance
    network-instance srl_default_instance {
        type default
        interface ethernet-1/1.2 {
            interface-ref {
                interface ethernet-1/1
                subinterface 2
            }
        }
        protocols {
            igmp {
                admin-state enable
                interface ethernet-1/1.2 {
                    version 3
                    query-interval 50
                    query-last-member-interval 20
                    query-response-interval 34
                }
            }
        }
```

Example: Configuring IGMP for an `ip-vrf` network-instance

The following example shows a configuration of IGMP for an `ip-vrf` type network instance (`srl_ipvrf_instance`). It includes enabling IGMP, configuring basic IGMP parameters, such as configuring IGMP on an interface, specifying the IGMP version for an interface, and setting up the query interval.

```
--{ +* candidate shared default }--[  ]--
# info with-context network-instance srl_ipvrf_instance
    network-instance srl_ipvrf_instance {
        type ip-vrf
        admin-state enable
        interface ethernet-1/1.4 {
            interface-ref {
                interface ethernet-1/1
                subinterface 4
            }
        }
        protocols {
            igmp {
                admin-state enable
                interface ethernet-1/1.4 {
                    version 3
                    query-interval 50
                    query-last-member-interval 20
                    query-response-interval 35
                }
            }
        }
    }
```

4.5.2 Configuring a static IGMP group membership

Procedure

You can configure static IGMP group memberships to test multicast forwarding without a receiver host. When IGMP static group membership is enabled, data is forwarded to an interface without receiving membership reports from host members.

When static IGMP group entries on point-to-point links that connect routers to a rendezvous point (RP) are configured, the static IGMP group entries do not generate join messages toward the RP. In an IGMP static group configuration, you can receive multicast traffic from multiple sources.

Example: Configuring a static IGMP group membership with a specific source IP

The following example configures a static IGMP group membership to ensure that the multicast traffic from sources `192.168.1.1` and `192.168.2.1` is forwarded to the specified multicast group range `239.1.1.1` to `239.1.1.20` without requiring dynamic IGMP join requests from the hosts.

```
--{ + candidate shared default }--[  ]--
# info with-context network-instance default protocols igmp interface ethernet-1/1.1
 static-membership-groups
```

```
    network-instance default {
        protocols {
            igmp {
               admin-state enable
                query-interval 50
                query-response-interval 34
                interface ethernet-1/1.1 {
                    static-membership-groups {
                        group-range 239.1.1.1 end 239.1.1.20 {
                            source 192.168.1.1 {
                            }
                            source 192.168.2.1 {
                            }
                        }
                    }
                }
            }
        }
    }
```

Example: Configuring a static IGMP group membership with starg entries

The following example configures a static IGMP group membership to ensure that the multicast traffic from any source is forwarded to the specified multicast group range `239.1.1.1` to `239.1.1.20` without requiring dynamic IGMP join requests from the hosts.

```
--{ + candidate shared default }--[  ]--
# info with-context network-instance default protocols igmp interface ethernet-1/1.1
 static-membership-groups
    network-instance default {
        protocols {
            igmp {
                interface ethernet-1/1.1 {
                    static-membership-groups {
                        group-range 239.1.1.1 end 239.1.1.20 {
                            starg
                        }
                    }
                }
            }
        }
    }
```

4.5.3 Configuring SSM translation

Procedure

You can configure static IGMP SSM mappings on the last-hop router to provide SSM translation for receiver hosts running IGMPv1 or IGMPv2.

Configuring the IGMP SSM mapping enables the multicast router to translate (*, G) information in IGMPv1 and IGMPv2 Report messages to (G, INCLUDE, (S1,S2..)) information, enabling SSM service for IGMPv1 and IGMPv2 hosts.

Example: Configuring SSM mapping with a specific source

In the following configuration, SSM mapping is set up to translate IGMPv1 or IGMPv2 join requests for multicast group range `239.1.1.1` to `239.1.1.20` into SSM join requests with a specific source. The SSM mapping ensures that any host attempting to join one of the multicast groups within this range is mapped to the source IP address 192.168.1.1. This means that even if a host does not explicitly specify the source when making a join request (as is the case with IGMPv1 and IGMPv2), the SSM mapping automatically associates the join request with the source 192.168.1.1.

```
--{ + candidate shared default }--[  ]--
# info with-context network-instance default protocols igmp ssm mappings
    network-instance default {
        protocols {
            igmp {
                ssm {
                    mappings {
                        group-range 239.1.1.1 end 239.1.1.20 {
                            source 192.168.1.1 {
                            }
                        }
                    }
                }
            }
        }
    }
```

Example: Configuring SSM mapping with multiple sources

In the following configuration, SSM mapping is set up to translate IGMPv1 or IGMPv2 join requests for multicast group 225.0.0.4 into SSM join requests for all the four mentioned sources (150.0.1.1, 150.0.1.2, 150.0.1.3, and 150.0.1.4), resulting in the receiver receiving four different multicast streams.

Note: SSM mapping with a combination of group range and different sources is also supported.

```
--{ + candidate shared default }--[  ]--
# info with-context network-instance default protocols igmp ssm mappings
    network-instance default {
        protocols {
            igmp {
                ssm {
                    mappings {
                        group-range 225.0.0.4 end 225.0.0.4 {
                            source 150.0.1.1 {
                            }
                            source 150.0.1.2 {
                            }
                            source 150.0.1.3 {
                            }
                            source 150.0.1.4 {
                            }
                        }
                    }
                }
            }
        }
    }
```

5 MLD

Multicast Listener Discovery (MLD) is the IPv6 version of IGMP and belongs to the Source Specific Multicast (SSM) service model. The purpose of MLD is to allow each IPv6 router to identify multicast listeners on its directly connected links and determine which multicast groups those neighboring nodes are interested in.

MLD is a sub-protocol of ICMPv6. MLD message types are a subset of the set of ICMPv6 messages, and MLD messages are identified in IPv6 packets by a preceding next header value of 58. All MLD messages are sent with a link-local IPv6 source address, a hop limit of 1, and an IPv6 router alert option in the hop-by-hop options header.

SR Linux supports the following MLD versions:

- MLDv1 (RFC 2710), derived from IGMPv2
- MLDv2 (RFC 3810), derived from IGMPv3

5.1 MLDv1

Similar to IGMPv2, MLDv1 reports only include the multicast group addresses that listeners are interested in and do not include the source addresses. To work with the PIM-SSM model, a similar SSM translation function is required when MLDv1 is used.

SSM translation allows an MLDv1 device to join an SSM multicast network through the router that provides such a translation capability. SSM translation per interface offers the flexibility of having the same (*, G) mapped to two different (S, G) on two different interfaces.

5.2 MLDv2

MLDv2 is backward compatible with MLDv1 and allows a node to report interest in listening to packets from a specific multicast group, either from particular source addresses or all sources except specified addresses.

5.3 MLD configuration

The routers use MLD to manage membership for a multicast session.

MLD is configured by enabling `mld admin-state` under the `network-instance protocols` context.

The `mld` model allows the configuration of MLD parameters in both the default and non-default network-instances, with the ability to apply specific configurations to individual interfaces.

When enabled, at least one interface must be specified in the MLD context, as MLD is an interface function. Traffic can only flow away from the router to an MLD interface and to and from a PIM interface. A router directly connected to a source must have PIM enabled on the interface to that source. The traffic travels in a network from the PIM interface to the PIM interface and arrives finally on an MLD enabled interface.

When MLD static group membership is enabled, data is forwarded to an interface without receiving membership reports from host members.

If static MLD group entries on point-to-point links that connect routers to a rendezvous point (RP) are configured, the static MLD group entries do not generate join messages toward the RP. When a host wants to receive multicast sessions, the host sends a join message for each multicast group it wants to join. A leave message may be sent for each multicast group with which it no longer needs to participate.

A multicast router keeps a list of multicast group memberships for each attached network and an interval timer for each membership. Hosts issue a multicast group membership report when they want to receive a multicast session. The reports are sent to all multicast routers.

5.3.1 MLD basic configuration

Procedure

MLD is configured by enabling `mld admin-state` under the `network-instance protocols` context.

Example: Configuring MLD for a default network-instance

The following example shows a configuration of MLD for a default network instance (`srl_default_instance`). It includes enabling MLD, configuring basic MLD parameters, such as configuring MLD on an interface, specifying the MLD version for an interface, and setting up the query interval.

```
--{ +* candidate shared default }--[  ]--
# info with-context network-instance srl_default_instance
    network-instance srl_default_instance {
        type default
        interface ethernet-1/1.2 {
            interface-ref {
                interface ethernet-1/1
                subinterface 2
            }
        }
        protocols {
            mld {
                admin-state enable
                interface ethernet-1/1.2 {
                    version 1
                    query-interval 125
                    query-last-member-interval 1
                    query-response-interval 10
                }
            }
        }
```

Example: Configuring MLD for an `ip-vrf` network-instance

The following example shows a configuration of MLD for an `ip-vrf` type network instance (`srl_ipvrf_instance`). It includes enabling MLD, configuring basic MLD parameters, such as configuring MLD on an interface, specifying the MLD version for an interface, setting up the query interval, and trace options.

```
--{ +* candidate shared default }--[  ]--
# info with-context network-instance srl_ipvrf_instance
```

```
    network-instance srl_ipvrf_instance {
        type ip-vrf
        admin-state enable
        interface ethernet-1/1.1 {
            interface-ref {
                interface ethernet-1/1
                subinterface 1
            }
        }
        protocols {
            mld {
                admin-state enable
                query-interval 125
                query-last-member-interval 1
                query-response-interval 10
                trace-options {
                    trace {
                        interface {
                            name ethernet-1/1.1
                        }
                    }
                }
                interface ethernet-1/1.1 {
                    version 1
                }
            }
        }
    }
```

5.3.2 Configuring a static MLD group membership

Procedure

Static MLD group memberships can be configured to test multicast forwarding without a receiver host. When MLD static group membership is enabled, data is forwarded to an interface without receiving membership reports from host members.

When static MLD group entries on point-to-point links that connect routers to a rendezvous point (RP) are configured, the static MLD group entries do not generate join messages toward the RP. In a static MLD group configuration, you can receive multicast traffic from multiple sources.

Example: Configuring a static MLD group membership with a specific source IP

The following example configures a static MLD group membership to ensure that the multicast traffic from source `2001:db8::1` is forwarded to the specified multicast group range `ff0e::db8:7` to `ff0e::db8:9` without requiring dynamic MLD join requests from the hosts.

```
--{ + candidate shared default }--[  ]--
# info with-context network-instance srl_default_instance protocols mld interface
 ethernet-1/1.1 static-membership-groups
    network-instance srl_default_instance {
        protocols {
            mld {
                interface ethernet-1/1.1 {
                    static-membership-groups {
                        group-range ff0e::db8:7 end ff0e::db8:9 {
                            source 2001:db8::1 {
                            }
                        }
                    }
                }
            }
        }
    }
```

Example: Configuring a static MLD multicast group with a specific address range and static (*,G) entry for the group

The following example configures a static MLD group membership to ensure that the multicast traffic from any source is forwarded to the specified multicast group range `ff0e::db8:7` to `ff0e::db8:9` without requiring dynamic MLD join requests from the hosts.

```
--{ + candidate shared default }--[  ]--
# info with-context network-instance srl_default_instance protocols mld interface
 ethernet-1/1.1 static-membership-groups
    network-instance srl_default_instance {
        protocols {
            mld {
                interface ethernet-1/1.1 {
                    static-membership-groups {
                        group-range ff0e::db8:7 end ff0e::db8:9 {
                            starg
                        }
                    }
                }
            }
        }
    }
```

5.3.3 Configuring SSM translation

Procedure

You can configure static MLD SSM mappings on the last-hop router to provide SSM support for receiver hosts that are running MLDv1. Configuring the IGMP SSM mapping enables the multicast router to translate (*, G) information in MLDv1 Report messages to (G, INCLUDE, (S1, S2..)) information, enabling SSM service for MLDv1 hosts.

Example: Configuring SSM mapping

In the following configuration, SSM mapping is set up to translate MLDv1 join requests for multicast group range `ff0e::db8:7` to `ff0e::db8:9` into SSM join requests with specific sources. The SSM mapping ensures that any host attempting to join one of the multicast groups within this range is mapped to the source IP address `2001:db8::1`. This means that even if a host does not explicitly specify the source when making a join request (as is the case with MLDv1), the SSM mapping automatically associates the join request with the source `2001:db8::1`.

```
--{ + candidate shared default }--[  ]--
# info with-context network-instance srl_default_instance protocols mld interface
 ethernet-1/1.1 ssm mappings
        protocols {
            mld {
                interface ethernet-1/1.1 {
                    ssm {
                        mappings {
                            group-range ff0e::db8:7 end ff0e::db8:9 {
                                source 2001:db8::1 {
                                }
                            }
                        }
                    }
                }
            }
        }
    }
```

6 Protocol Independent Multicast (PIM)

PIM leverages the unicast routing protocols such as OSPF, IS-IS, BGP, and static routes to create the unicast routing table. Since PIM relies on unicast routing information to carry out the multicast forwarding function, it is effectively IP protocol independent. Unlike DVMRP, PIM does not send multicast routing table updates to its neighbors.

The following summarizes SR Linux support for PIM:

- PIM-SM (ASM model) on default and non-default network instances for both IPv4 and IPv6
- PIM-SM (SSM model) on default and non-default network instances for both IPv4 and IPv6
- PIM policy - IPv4 and IPv6

The logical path taken through the network from the source to multiple receivers is called the multicast distribution tree (MDT).

PIM has two modes of operation - Dense and Sparse mode

In PIM dense mode, the type of MDT used is called the source tree. The root of the tree is the source device, and data flows from the root (source) to the leaves through branches. This tree is also known as the shortest-path tree (SPT).

PIM sparse mode (PIM-SM) uses two types of multicast trees: the shared tree and the source tree.

The ASM mode of operation relies on the concept of a rendezvous point (RP), which is the point in the network where the source and receivers meet to establish the multicast flow. The shared tree is a tree rooted at the RP. Multicast transmission starts with the source sending packets to the RP and from there on, the shared tree to the receivers. ASM is represented by `(*,G)` notation. The `*` (wildcard) indicates that the multicast group can receive traffic from any source, and `G` represents the multicast group address.

In the SSM mode of operation, the RP is omitted, and the source tree is established directly. SSM is represented by `(S,G)` notation. The `S` indicates that the multicast group receives traffic from a specific source, and `G` represents the multicast group address.

6.1 PIM sparse mode (PIM-SM)

PIM-SM only forwards data to network segments with active receivers that have explicitly requested the multicast group. PIM-SM uses the unicast routing table to perform the Reverse Path Forwarding (RPF) check function instead of building up a completely independent multicast routing table.

Depending on the configuration options, the traffic can remain on the shared tree or switch over to an optimized source distribution tree. As multicast traffic starts to flow down the shared tree, routers along the path determine if there is a better path to the source. If a more direct path exists, then the router closest to the receiver sends a join message toward the source and then reroutes the traffic along this path.

As stated above, PIM-SM relies on an underlying topology-gathering protocol to populate a routing table with routes. This routing table is called the Multicast Routing Information Base (MRIB). The routes in this table can be taken directly from the unicast routing table, or they can be different and provided by a separate routing protocol such as Multiprotocol BGP (MBGP) which is used for Next Generation Multicast VPN (NG-MVPN). Regardless of how it is created, the primary role of the MRIB in the PIM-SM protocol is to provide the next hop router along a multicast-capable path to each destination subnet. The MRIB is used to determine the next hop neighbor to whom any PIM join/prune message is sent. Data flows along the reverse path of the join messages. Thus, in contrast to the unicast RIB that specifies the next hop that a data packet would take to get to some subnet, the MRIB gives reverse-path information and indicates the path that a multicast data packet would take from its origin subnet to the router that has the MRIB.

Note:
For the PIM protocol to function correctly, multicast signaling (PIM) packets need to be received by the CPM CPU. Therefore, CPM filters and management access filters must be configured to allow forwarding of PIM packets. PIM signaling needs to signal the specific <S, G> from the host to RP or the source of the tree. Using policies, specific <S, G>s can be disallowed to join the tree. If a <S, G> is disallowed to join the tree, any multicast stream from that <S, G> is dropped on the datapath.

6.1.1 PIM-SM functions

PIM-SM functions have three phases.

Phase one

In this phase, a multicast receiver expresses its interest in receiving traffic destined for a multicast group. Typically, it does this using IGMP, but other mechanisms may also serve this purpose. One of the local routers of the receiver is elected as the Designated Router (DR) for that subnet. When the expression of interest is received, the DR sends a PIM join message toward the Rendezvous Point (RP) for that multicast group. This join message is known as a (*, G) join because it joins group G for all sources to that group. The (*, G) join travels hop-by-hop toward the RP for the group, and in each router it passes through, the multicast tree state for group G is instantiated.

Eventually, the (*, G) join either reaches the RP or reaches a router that already has the (*, G) join state for that group. When many receivers join the group, their join messages converge on the RP and form a distribution tree for group G that is rooted at the RP. The distribution tree is called the RP tree or the shared tree because it is shared by all sources transmitting to the same multicast group. Join messages are re-sent periodically as long as the receiver remains in the group. When all receivers on a leaf network leave the group, the DR sends a PIM (*, G) prune message toward the RP for that multicast group. However, if the prune message is not sent for any reason, the state eventually times out.

A multicast data sender starts sending data destined for a multicast group. The local router of the sender (the DR) takes these data packets, unicast-encapsulates them, and sends them directly to the RP. The RP receives these encapsulated data packets, removes the encapsulation, and forwards them to the shared tree. The packets then follow the (*, G) multicast tree state in the routers on the RP tree, and are replicated wherever the RP tree branches, and eventually reach all the receivers for that multicast group. The process of encapsulating data packets to the RP is called registering, and the encapsulation packets are known as PIM register packets.

At the end of phase one, multicast traffic flows encapsulated to the RP and then natively over the RP tree to the multicast receivers.

Phase two

In this phase, register-encapsulation of data packets is performed. However, register-encapsulation of data packets is inefficient for the following reasons:

- Encapsulation and de-encapsulation can be resource-intensive operations for a router to perform, depending on whether the router has appropriate hardware for the tasks.

- Traveling to the RP and then back down the shared tree can cause the packets to travel a relatively long distance to reach receivers that are close to the sender. For some applications, increased latency is unwanted.

Although register-encapsulation can continue indefinitely, for the previous reasons, the RP normally switches to native forwarding. To do this, when the RP receives a register-encapsulated data packet from source S on group G, it normally initiates an (S, G) source-specific join toward S. This join message travels hop-by-hop toward S, instantiating an (S, G) multicast tree state in the routers along the path.

The (S, G) multicast tree state is used only to forward packets for group G if those packets come from source S. Eventually, the join message reaches the S subnet or a router that already has the (S, G) multicast tree state, and packets from S start to flow following the (S, G) tree state toward the RP. These data packets can also reach routers with a (*, G) state along the path toward the RP, and if this occurs, they take a shortcut to the RP tree at this point.

While the RP is in the process of joining the source-specific tree for S, the data packets continue being encapsulated and sent to the RP. When packets from S also start to arrive natively at the RP, the RP receives two copies of each of these packets. At this point, the RP starts to discard the encapsulated copy of these packets and sends a register-stop message back to the DR of S to prevent the DR from unnecessarily encapsulating the packets. At the end of phase two, traffic is flowing natively from S along a source-specific tree to the RP and from there along the shared tree to the receivers. Where the two trees intersect, traffic can transfer from the shared RP tree to the shorter source tree.

Note:
A sender can start sending before or after a receiver joins the group; therefore, phase two may occur before the shared tree to the receiver is built.

Phase three

In this phase, the RP joins back toward the source using the shortest path tree (SPT). Although having the RP join back toward the source removes the encapsulation overhead, it does not completely optimize the forwarding paths. For many receivers, the route via the RP can involve a significant detour when compared with the shortest path from the source to the receiver.

To obtain lower latencies, a router on the LAN of the receiver, typically the DR, may optionally initiate a transfer from the shared tree to a source-specific SPT. To do this, it issues an (S, G) join toward S. This instantiates the (S, G) state in the routers along the path to S. Eventually, this join either reaches the S subnet or reaches a router that already has the (S, G) state. When this happens, data packets from S flow following the (S, G) state until they reach the receiver.

At this point, the receiver (or a router upstream of the receiver) receives two copies of the data—one from the SPT and one from the RP tree. When the first traffic starts to arrive from the SPT, the DR or upstream router starts to drop the packets for G from S that arrive via the RP tree. In addition, it sends an (S, G) prune message toward the RP. The prune message travels hop-by-hop, instantiating an (S, G) state along the path toward the RP, indicating that traffic from S for G should not be forwarded in this direction. The prune message is propagated until it reaches the RP or a router that still needs the traffic from S for other receivers.

By now, the receiver is receiving traffic from S along the SPT between the receiver and S. In addition, the RP is receiving the traffic from S, but this traffic is no longer reaching the receiver along the RP tree. As far as the receiver is concerned, this is the final distribution tree.

6.1.2 Differences between PIM-SM in ASM and SSM models

Table 5: Differences between PIM-SM in ASM and SSM models

Features	PIM-SM in ASM model	PIM-SM in SSM model
Source discovery	Uses RP for source discovery	Does not require source discovery and only supports a single source for a specific multicast stream
Rendezvous Point (RP)	Requires an RP for initial traffic distribution and source discovery	Does not require RP. The traffic is source-specific
Join process	Receivers join a multicast group and expect traffic from any source	Receivers join a multicast group and specify the source they want to receive traffic from
Traffic distribution	Initially uses the RP to forward traffic, then switches to the SPT	Directly uses the SPT from the start without using an RP

6.1.3 Encapsulating data packets in the register tunnel

Conceptually, the register tunnel is an interface with a smaller Maximum Transmission Unit (MTU) than the underlying IP interface toward the RP. IP fragmentation on packets forwarded on the register tunnel is performed based on this smaller MTU. The encapsulating DR can perform path-MTU discovery to the RP to determine the effective MTU of the tunnel. This smaller MTU considers both the outer IP header and the PIM register header overhead.

6.1.4 PIM-SM routing policies

Multicast traffic can be restricted from specific source addresses by creating routing policies.

SR Linux route policy match criteria for PIM-SM can specify the following:

- multicast group address embedded in the join and prune message
- multicast source address embedded in the join and prune message
- multicast group address embedded in the register message
- multicast source address embedded in the register message

When a multicast policy is applied to IGMP, MLD, or PIM using the command `routing-policy+ policy* [name] statement** [name] match multicast <group-address,source-address>`, the multicast parameters in the policy statement are compared to the source and group (S, G) or (*, G) in the incoming join or prune messages for the protocol. If a match is found, the join or prune message is accepted or denied based on the policy action. The same policy can be applied to multiple protocols.

6.1.5 RPF checks

Multicast implements a Reverse Path Forwarding (RPF) check. RPF checks the path that multicast packets take between their sources and the destinations to prevent loops. Multicast requires that an incoming interface be the outgoing interface used by unicast routing to reach the source of the multicast packet. RPF forwards a multicast packet only if it is received on an interface that is used by the router to route to the source.

If the forwarding paths are modified because of routing topology changes, any dynamic filters that may have been applied must be re-evaluated. If filters are removed, the associated alarms are also cleared.

6.1.6 Distributing PIM joins over multiple ECMP paths

The per bandwidth or round-robin method is commonly used for multicast load-balancing. However, the interface in an ECMP set can also be used for a specific channel to be predictable without knowledge of other channels that use the ECMP set.

When a link in the ECMP set is removed, multicast streams that use this link are redistributed over the remaining ECMP links using the same hash algorithm. When a link is added to the ECMP set, new joins may be allocated to the new link based on the hash algorithm. Existing multicast streams using the other ECMP links stay on those links until they are pruned, unless the `rebalance` parameter under `network-instance* [name] protocols pim!+ ecmp-hashing!` context is enabled.

The default is not enabled, which means that the use of multiple ECMP paths is controlled through the existing implementation and the enabling of `ecmp-balance` option under `network-instance* [name] protocols pim!+` context.

Note:
You cannot enable the `ecmp-balance` option when the MC ECMP hashing option `ecmp-hashing` is configured in the same context.

To achieve distribution of streams across the ECMP links, the hashing steps are as follows:

1. For a specific (S,G) get all possible next hops.
2. Sort these next hops based on next hop address.
3. XOR S and G addresses.
4. Hash the XORed address over the number of PIM next hops.
5. Use the hash value obtained in step 4, and set that element in the sorted list that was obtained in step 2 as the preferred next hop
6. If this element is not available or is not a PIM next hop (PIM neighbor), choose the next available next hop.

6.2 IPv6 PIM models

IPv6 multicast enables multicast applications over native IPv6 networks.

There are two service models:

- Any Source Multicast (ASM)

- Source Specific Multicast (ASM)

These models work in conjunction with PIM-SM and MLD to delivery multicast in IPv6 network.SSM does not require source discovery and only supports single source for a specific multicast stream. As a result, SSM is easier to operate in a large scale deployment that uses the one-to-many service model.

6.2.1 PIM SSM

The SSM model supports the IPv6 address family and allows for selecting the Routing Table Manager (RTM), whether unicast RTM, multicast RTM, or both. OSPFv3, IS-IS, and static routes have extensions that enable route submissions to the IPv6 multicast RTM.

6.2.2 PIM ASM

The ASM model supports the IPv6 address family. All PIM ASM-related functions, such as RP and SPT, support IPv4 and IPv6 address-families.

6.3 Forwarding rate measurement for PIM groups (7730 SXR)

The forwarding rate represents the current rate at which multicast traffic is forwarded for a specific PIM group. This rate is typically measured in bits per second (bps).

Note: On 7220 IXR-D*x* platforms, the forwarding rate for a specific PIM group cannot be measured.

6.4 Basic PIM configuration

The PIM protocol is operational after at least one interface is assigned. After an interface is enabled for PIM, it is known as a PIM interface. When created, the PIM interface can be configured with PIM parameters in addition to the standard parameters for the interface. When PIM is operational, data is forwarded to network segments with active host receivers that have explicitly requested the multicast group.

To configure PIM, perform the following tasks:

1. Enable PIM (required)
2. Add interfaces so the protocol establishes adjacencies with the neighboring routers (required)
3. Configure a way to calculate group-to-RP mapping (required) by static group-to-RP mapping
4. Enable unicast routing protocols to learn routes toward the RP/source for reverse path forwarding (required)
5. Add SSM ranges (optional)
6. Change hello interval (optional)
7. Configure route policies

6.4.1 Configuring IPv4 PIM-SM in the ASM model

Procedure

In PIM-ASM, routers forward multicast traffic only to the receivers that explicitly request it via an RP. Receivers use IGMPv1 or IGMPv2 to join a multicast group without specifying the source. The RP initially forwards the multicast traffic from any source to the receivers through a Shared Tree (RPT). Later, the network may switch to a Source-Specific Tree (SPT).

Example: Configuring PIM-SM with static RP using the ASM model

The following configuration enables PIM on interface `ethernet-1/1.1` and defines a static RP with the IP address `10.10.10.1`. The RP manages the multicast traffic for the groups with the prefixes `239.24.24.24/32` and `239.24.28.24/32`.

```
--{ + candidate shared default }--[  ]--
# info with-context network-instance srl_test_instance protocols pim
    network-instance srl_test_instance {
        protocols {
            pim {
                interface ethernet-1/1.1 {
                    admin-state enable
                }
                rendezvous-points {
                    static {
                        rendezvous-point 10.10.10.1 {
                            group 239.24.24.24/32 {
                            }
                            group 239.24.28.24/32 {
                            }
                        }
                    }
                }
            }
        }
    }
```

6.4.2 Configuring IPv4 PIM-SM in the SSM model

Procedure

In PIM-SM, routers forward multicast traffic only to the receivers that explicitly request it through SSM group membership using IGMPv3. In this model, the receivers specify both the multicast group and the source they want to receive traffic from; therefore, no RP is needed. The source address is included in the IGMP report messages that are passed to PIM routers as PIM join messages. The multicast traffic is directly delivered from the specified source to the receivers using a source-specific tree (SPT).

Example: Configuring PIM-SM using the SSM model

The following configuration enables PIM on interface `ethernet-1/1.1` and uses IGMPv3 to manage multicast group memberships. The SSM group address range is, by default, `232.0.0.0/8`.

```
--{ + candidate shared default }--[  ]--
# info with-context network-instance srl_test_instance protocols pim ssm
    network-instance srl_test_instance {
        protocols {
            pim {
                ssm {
                    ssm-ranges {
                        group-range 233.0.0.0/8 {
                        }
                    }
                }
            }
        }
```

6.4.3 Configuring IPv6 PIM-SM in the ASM model

Procedure

The following configuration enables PIM on interface `ethernet-1/1.1` and defines a static RP with the IP address `2001:db8:1:1::1`. The RP manages the multicast traffic for the group with the prefix `ff05::/16`.

Example

```
--{ + candidate shared default }--[  ]--
# info with-context network-instance srl_ipv6_test_instance protocols pim
    network-instance srl_ipv6_test_instance {
        protocols {
            pim {
                interface ethernet-1/1.1 {
                    admin-state enable
                }
                rendezvous-points {
                    static {
                        rendezvous-point 2001:db8:1:1::1 {
                            group ff05::/16 {
                            }
                        }
                    }
                }
            }
        }
    }
```

6.4.4 Configuring IPv6 PIM-SM in the SSM model

Procedure

The following configuration enables PIM on interface `ethernet-1/1.1` and uses IGMPv3 to manage multicast group memberships. By default, the SSM group address range is `FF3E::/32`.

Example

```
--{ + candidate shared default }--[  ]--
# info with-context network-instance srl_test_instance protocols pim ssm
    network-instance srl_test_instance {
        protocols {
            pim {
                ssm {
                    ssm-ranges {
                        group-range ff3e:1:0:0::/64 {
                        }
                    }
                }
            }
        }
```

6.4.5 Configuring PIM hello messages

Procedure

By default, the neighboring PIM routers periodically send PIM Hello messages to each other every 30 seconds. You can configure the PIM Hello interval (`hello-interval`) to adjust the frequency of the hello message. The `hold-time` specifies how long the router should maintain information about a neighbor after receiving its PIM hello message
When a router receives a PIM hello message, it stores the neighbor's IP address and DR priority. A designated router (DR) election is performed on all multi-access networks. The DR is the device with the highest configured DR priority on the broadcast domain. If priorities are equal, the DR is the device with the highest IP address on the broadcast domain.

Example

In the following configuration example, the PIM `hello-interval` and `hello-multiplier` are set to 65, defining how often hello messages are sent and their expiration time.

```
--{ + candidate shared default }--[  ]--
# info with-context network-instance default_nw_instance protocols pim
  network-instance default_nw_instance {
        protocols {
            pim {
                admin-state enable
                interface ethernet-1/1.1 {
                    hello-interval 65
                    hello-multiplier 65
                }
                rendezvous-points {
                    static {
                        rendezvous-point 10.10.10.1 {
                            group 239.24.24.24/32 {
                            }
                        }
                    }
                }
            }
        }
    }
```

6.4.6 Configuring PIM join policies

Procedure

Configuring PIM join policies establishes a routing policy that enables the router to manage multicast traffic effectively by allowing only specific combinations of source and group addresses. Before configuring a multicast routing policy, you must configure prefix sets for the group and source addresses.

Example: Configuring prefix set for group address

The following configuration example defines a group prefix (`srl_group_prefix-set`) that can be used as a matched against `group-address` configured in the policy statement (`allow-group`).

```
--{ * candidate shared default }--[  ]--
# info with-context routing-policy prefix-set srl_group_prefix-set
    routing-policy {
        prefix-set srl_group_prefix-set {
            prefix 226.0.0.0/21 mask-length-range 21..24 {
            }
        }
    }
```

Example: Configuring prefix set for source address

The following configuration example defines a group prefix (`srl_source_prefix-set`) that can be used as a matched against `source-address` configured in the policy statement (`allow-source`).

```
--{ * candidate shared default }--[  ]--
# info with-context routing-policy prefix-set srl_source_prefix-set
    routing-policy {
        prefix-set srl_source_prefix-set {
            prefix 192.0.0.0/21 mask-length-range 21..25 {
            }
        }
    }
```

Example: Configuring multicast routing policy

In the following configuration example, the policy (`multicast_new_policy`) allows traffic that matches the specified group and source prefixes.

```
--{ * candidate shared default }--[  ]--
# info with-context routing-policy policy multicast_new_policy
    routing-policy {
        policy multicast_routing_policy {
            default-action {
                policy-result reject
            }
            statement allow-group {
                match {
                    multicast {
                        group-address {
                            prefix-set srl_group_prefix-set
                        }
                    }
                }
            }
            statement allow-source {
                match {
                    multicast {
                        source-address {
                            prefix-set srl_source_prefix-set
                        }
                    }
                }
            }
        }
    }
```

6.4.7 Configuring the switchover to SPT (7730 SXR)

Procedure

Note: Switching to SPT mode is not supported on the 7220 IXR-D*x* platforms.

By default, the last hop router or the receiver DR initiates an SPT switchover toward the source after receiving the first copy of multicast data packets.

Use the `spt-switchover` parameter in the `network-instance* [name] protocols pim!+ spt-switchover group* [prefix] threshold?` context, to define the threshold for switching from a shared tree (RP) to a shortest path tree (SPT). Configure the `spt-switchover` parameter with a specific group prefix to define the switchover threshold on a per-group basis.

When the SPT threshold is reached, and the traffic switches over to SPT, a PIM (*, G) prune message is sent up the shared tree towards RP. This ensures multicast data is exclusively sent over the SPT.
To prevent the DR from triggering an SPT switchover, set the `spt-switchover` parameter to `infinity`.

Example

In the following configuration, the SPT switchover threshold is set for the multicast group prefix `239.255.0.0/16` with a threshold of `100000` kbps. When multicast traffic directed to any address within the specified group prefix exceeds `100000` kbps; the router transitions from the shared path tree to SPT.

```
--{ + candidate shared default }--[  ]--
# info with-context network-instance srl_test_intance protocols pim
    network-instance srl_test_intance {
        protocols {
            pim {
                admin-state enable
                interface ethernet-1/1.1 {
                }
                spt-switchover {
                    group 239.255.0.0/16 {
                        threshold 100000
                    }
                }
            }
        }
    }
```

7 Multicast Source Discovery Protocol (MSDP)

Note: Multicast Source Discovery Protocol (MSDP) is currently supported only on the 7250 IXR Gen 2/2c+ platforms and exclusively within the default network instance..

The Multicast Source Discovery Protocol (MSDP), an inter-domain multicast protocol, enables Rendezvous Points (RPs) in different PIM-SM domains to discover active multicast source information among them. RP is the point in the PIM-SM domain where the source and receivers meet to establish the multicast flow.

MSDP uses TCP port 639 for peering. Without MSDP, RPs in one domain would be unaware of active multicast sources in other domains, preventing the establishment of multicast peer paths across autonomous boundaries.

MSDP is required in PIM-SM Any-Source Multicast (ASM) domain mode, but is not necessary for Source-Specific Multicast (SSM) domain mode. Unlike ASM, SSM does not depend on RPs or inter-domain source discovery. Instead, the network can directly construct multicast forwarding trees based on the explicitly specified source and group.

Service providers whose multicast networks span multiple autonomous systems (AS) use MSDP to establish peering relationships for multicast exchange.

- **MSDP speaker:** the router within a PIM-SM domain that establishes an MSDP peering session with MSDP peers in other domains
- **MSDP peer:** the router that establishes a peering session with the MSDB speaker
- **MSDP peering session:**the TCP connection used for exchanging MSDP control information between peers.
- **Source-Active (SA) message:** When an RP in a PIM-SM domain first discovers a new sender (source), for example, through PIM register messages, it generates a source-active (SA) message and forwards it to its MSDP peers. RPs that originate SA messages continue to do so periodically as long as the source is actively sending data. Although the SA message contains several fields, the most significant are:
  - source address (*S*) of the data source
  - group address (*G*) to which the source sends
  - IP address of the originating RP
- **MSDP peer groups:** MSDP peer groups are typically created when multiple peers have a set of common operational parameters. Group parameters that are not specifically configured are inherited from the global level.
- **MSDP mesh groups:** MSDP mesh groups are used to reduce SA flooding primarily in intra-domain configurations. When a number of speakers in an MSDP domain are fully meshed, they can be configured as a mesh group. The originator of the SA message forwards the message to all members of the mesh group. Because of this, forwarding the SA between non-originating members of the mesh group is not necessary.

7.1 MSDP procedure

The MSDB operational procedure is referred to as the flood-and-join procedure because the SA messages are flooded across MSDP peers, and only RPs with interested group members join the multicast distribution tree, while others ignore the SA message.

The sequence of events in the MSDB operational procedure is as follows:

1. When an RP in a PIM-SM domain first discovers a new sender (source), for example, through PIM register messages, it generates a source-active (SA) message and forwards it to its MSDP peers.
2. Each MSDP peer receives and forwards the SA message away from the RP address in a peer-RPF flooding fashion. The peer-RPF flooding applies to forwarding SA messages. The Multicast Routing Information Base (MRIB) is examined to determine which peer toward the originating RP of the SA message is selected. Such a peer is called an RPF peer. If the MSDP peer receives the SA message from a non-RPF peer toward the originating RP, it drops the message. Otherwise, it forwards the message to all its MSDP peers (except the one from which it received the SA message). For additional rules, see Peer-RPF check.

   Note: In regular multicast RPF checks, the packet's source address is compared against the interface upon which the packet was received. In PIM-SM with MSDP implementation, the peer-RPF check compares the RP address carried in the SA message against the MSDP peer from which the message was received.

3. An MSDP peer that is also an RP for its own domain receives a new SA message, and it determines if any group members within the domain are interested in any group described by an (S,G) entry within the SA message.
4. The RP (the receiving RP) checks for a (*,G) entry with a non-empty outgoing interface list. This implies that some system in the domain is interested in the group.
5. The RP (the receiving RP) triggers an (S,G) join event toward the data source as if a join or prune message addressed to the RP was received. This sets up a branch of the source tree for this domain.
6. Subsequent data packets arrive at the RP by this tree branch and are forwarded down the shared tree inside the domain.
7. If leaf routers choose to join the source-tree, they have the option to do so according to existing PIM-SM conventions. If an RP in a domain receives a PIM join message for a new group G, the RP must trigger an (S,G) join event for each active (S,G) for that group in its SA cache.

7.1.1 Peer-RPF check

Unlike the regular multicast RPF checks, the peer-RPF check stops SA messages from looping.
An MSDP router validates SA messages originated from other routers in a deterministic fashion. When the router receives an SA message, it applies a set of rules to validate the SA message, and the first rule that applies determines the peer-RPF neighbor. All SA messages from other routers are rejected. The rules used to validate SA messages originating at Router S received at Router R from Router N are as follows:

1. If the Router N and Router S are the same, the message is originated by a direct peer-RPF neighbor and is accepted.
2. If Router N is a configured peer or a member of the Router R mesh group, its SA messages are accepted.

3. If Router N is the Broader Gateway Protocol (BGP) next hop for the active multicast RPF route to Router S, then Router N is the peer-RPF neighbor, and its Source Active (SA) messages are accepted
4. If Router N is an external BGP peer of Router R and the last autonomous system (AS) number in the BGP AS-path to Router S is the same as the AS number of Router N, Router N is the peer-RPF neighbor, and its SA messages are accepted.
5. If Router N uses the same next hop as the next hop to Router S, Router N is the peer-RPF neighbor, and its SA messages are accepted.
6. If Router N does not fit any of the preceding rules, it is not a peer-RPF neighbor, and its SA messages are rejected.

When a peer is configured as a default peer, all SA messages received from the peer are accepted without performing the preceding peer-RPF check.

7.2 MSDP peer scenarios

RFC 4611, *Multicast Source Discovery Protocol (MSDP) Deployment Scenarios*, describes how protocols interact to deliver intra- and inter-domain ASM services.

Inter-domain peering:

- peering between PIM border routers (single-hop peering)
- peering between non-border routers (multi-hop peering)
- MSDP peering without BGP
- MSDP peering between mesh groups
- MSDP peering at a multicast exchange

Intra-domain peering:

- peering between routers configured for both MSDP and MBGP
- MSDP peer is not a BGP peer (meaning, no BGP peer)

7.3 MSDP configurations

MSDP is configured by enabling `msdp.admin-state` within the `network-instance.protocols` context.

The MSDP (`msdp`) protocol allows the MSDP parameter configuration **only** in the `default` network instance type.

7.3.1 Configuring MSDP for a default network-instance

Procedure

The minimum MSDP configuration requires enabling the **msdp.admin-state** parameter under the `network-instance protocols` context, configuring at least one MSDP peer to establish an SA relationship, and setting a local address.

Example

In the following configuration, MSDP is enabled with a standard group (`srl_test`) containing two peers (`10.10.10.104` and `10.10.10.106`) and an additional standalone peer (`10.20.1.1`) set to local address `10.20.1.6`.

```
--{ + candidate shared default }--[  ]--
# info with-context network-instance default protocols msdp
    network-instance default {
        protocols {
            msdp {
                admin-state enable
                group srl_test {
                    active-source-limit 50000
                    mode standard
                    receive-message-rate {
                        rate 100
                        time 300
                        threshold 5000
                    }
                    peer 10.10.10.104 {
                    }
                    peer 10.10.10.106 {
                    }
                }
                peer 10.20.1.1 {
                    local-address 10.20.1.6
                }
            }
        }
    }
```

Customer document and product support

Customer documentation

Customer documentation welcome page

Technical support

Product support portal

Documentation feedback

Customer documentation feedback