

Chapter 28

PIM Overview

Protocol Independent Multicast (PIM) is used for efficiently routing to multicast groups that might span wide-area and interdomain internetworks. It is called “protocol independent” because it does not depend on a particular unicast routing protocol. The JUNOS software supports sparse mode, dense mode, and sparse-dense mode.

For information about standards supported for PIM, see “IP Multicast Standards” on page 20.

PIM Modes

Because the mode you choose determines the PIM configuration properties, you first must decide whether PIM operates in sparse, dense, or sparse-dense mode in your network. Each mode has distinct operating advantages in different network environments.

In sparse mode, routers must join and leave multicast groups explicitly. Upstream routers do not forward multicast traffic to a router unless it has sent an explicit request (by means of a join message) to the rendezvous point (RP) router to receive this traffic. The RP serves as the root of the shared multicast delivery tree and is responsible for forwarding multicast data from different sources to the receivers.

Sparse mode is well suited to the Internet, where frequent interdomain joins and prunes are common.

Unlike sparse mode, in which data is forwarded only to routers sending an explicit PIM join request, dense mode implements a *flood-and-prune* mechanism, similar to the Distance Vector Multicast Routing Protocol (DVMRP). In dense mode, a router receives the multicast data on the incoming interface, then forwards the traffic to the outgoing interface list. Flooding occurs periodically, and is used to refresh state information, such as the source IP address and multicast group pair. If the router has no interested receivers for the data, and the outgoing interface list becomes empty, the router sends a PIM prune message upstream.

Dense mode works best in networks where few or no prunes occur. In such instances, dense mode is actually more efficient than sparse mode.

Sparse-dense mode, as the name implies, allows the interface to operate on a per-group basis in either sparse or dense mode. A group specified as “dense” is not mapped to an RP. Instead, data packets destined for that group are forwarded by means of PIM dense mode rules. A group specified as “sparse” is mapped to an RP, and data packets are forwarded by means of PIM sparse-mode rules.

Sparse-dense mode is useful in networks implementing auto-RP for PIM sparse mode.

For more information about how the PIM modes operate, see:

- PIM Sparse Mode on page 186
- PIM Dense Mode on page 203
- PIM Sparse-Dense Mode on page 204
- RP Mapping with Anycast RP on page 205
- Multicast over Layer 3 VPNs on page 205
- Tunnel Services PICs and Multicast on page 206
- Filtering Multicast Messages on page 207
- Embedded RP for IPv6 Multicast on page 209

For more information about mode-dependent configurations, see:

- Configuring PIM Dense Mode Properties on page 220
- Configuring PIM Sparse Mode Properties on page 221
- Configuring Sparse-Dense Mode Properties on page 240

PIM Sparse Mode

A PIM sparse-mode domain uses reverse-path forwarding (RPF) to create a path from a data source to the receiver requesting the data. When a receiver issues an explicit join request, an RPF check is triggered. A (*,G) PIM join message is sent toward the RP from the receiver’s designated router (DR). (By definition, this message is actually called a join/prune message, but for clarity in this description, it is called either join or prune, depending on its context.) The join message is multicast hop by hop upstream to the ALL-PIM-ROUTERS group (224.0.0.13) by means of each router’s RPF interface until it reaches the RP. The RP router receives the (*,G) PIM join message and adds the interface on which it was received to the OIL of the rendezvous-point tree (RPT) forwarding state entry. This builds the RPT connecting the receiver with the RP. The RPT remains in effect, even if no active sources generate traffic.



NOTE: State—the (*,G) or (S,G) entries—is the information used for forwarding unicast or multicast packets. S is the source IP address, G is the multicast group address, and * represents any source sending to group G. Routers keep track of the multicast forwarding state for the incoming and outgoing interfaces for each group.

When a source becomes active, the source's DR encapsulates multicast data packets into a PIM register message and sends them by means of unicast to the RP router.

If the RP router has interested receivers in the PIM sparse-mode domain, it sends a PIM join message toward the source to build a shortest-path tree (SPT) back to the source. The source sends multicast packets out on the LAN, and the source's DR encapsulates the packets in a PIM register message and forwards it toward the RP router by means of unicast. The RP router receives PIM register messages back from the source, and thus adds a new source to the distribution tree, keeping track of sources in a PIM table. Once an RP router receives packets natively (with S,G), it sends a register stop message to stop receiving the register messages by means of unicast.

In actual application, many receivers with multiple SPTs are involved in a multicast traffic flow. To simply illustrate the process, we track the multicast traffic from the RP router to one receiver. In such a case, the RP router begins sending multicast packets down the RPT toward the receiver's DR for delivery to the interested receivers. When the receiver's DR gets the first packet from the RPT, the DR sends a PIM join message toward the source's DR to start building an SPT back to the source. When the source's DR receives the PIM join message from the receiver's DR, it starts sending traffic down all SPTs. When the first multicast packet is received by the receiver's DR, the receiver's DR sends a PIM prune message to the RP router to stop duplicate packets being sent through the RPT. In turn, the RP router stops sending multicast packets to the receiver's DR, and sends a PIM prune message for this source over the RPT toward the source DR to halt multicast packet delivery to the RP router from that particular source.

If the RP router receives a PIM register message from an active source, but has no interested receivers in the PIM sparse-mode domain, it still adds the active source into the PIM table. However, after adding the active source into the PIM table, the RP router sends a register stop message. The RP router knows of the active source's existence and no longer needs to receive advertisement of the source (which utilizes resources).

This section contains more information about the routers and PIM sparse-mode functions briefly described above:

- Designated Router on page 188
- Rendezvous Point on page 189
- RP Mapping Options on page 189
- Building an RPT Between RP and Receivers on page 192
- PIM Sparse-Mode Source Registration on page 193
- PIM Sparse-Mode SPT Cutover on page 196
- PIM SSM on page 200

Designated Router

In a PIM sparse-mode domain, there are two types of designated routers to consider:

- The receiver's DR sends PIM join and PIM prune messages from the receiver network toward the RP.
- The source's DR sends PIM register messages from the source network to the RP.

Regardless of whether it is the receiver's DR or the source's DR, a DR is selected from other routers in a network by the exchange of IP addresses. Neighboring PIM sparse-mode routers multicast periodic PIM hello messages to each other every 30 seconds (the default). On receipt of a PIM hello message, a router stores the IP address and priority for that neighbor. If the DR priorities match, the router with the highest IP address is selected as the DR.

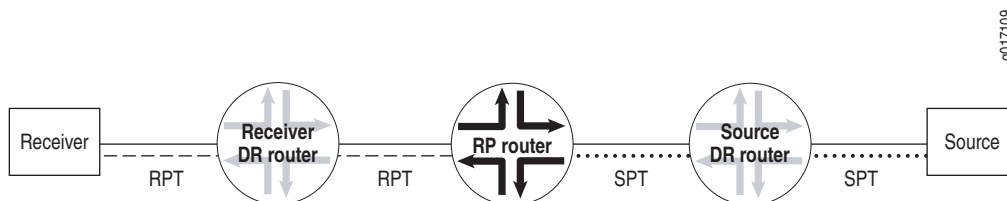
If a DR fails, a new one is selected using the same process of comparing IP addresses.

Rendezvous Point

The RP router serves as the information exchange point for the other routers. All routers in a PIM domain must provide mapping to an RP router. It is the only router that needs to know the active sources for a domain—the other routers just need to know how to get to the RP. In this way, the RP matches receivers with sources.

The RP router is downstream from the source and forms one end of the SPT. As shown in Figure 14, the RP router is upstream from the receiver and thus forms one end of the RPT.

Figure 14: The RP as Part of the RPT and SPT



The benefit of using the RP as the information exchange point is that it reduces the amount of state in non-RP routers. No network flooding is required to provide non-RP routers information about active sources.

RP Mapping Options

RPs can be learned by one of the following mechanisms:

- Static Configuration on page 189
- Anycast RP on page 190
- Auto-RP on page 190
- Bootstrap Router on page 191

Static Configuration

You can configure a static RP configuration that is very similar to static routes. A static configuration has the benefit of operating in PIM version 1 or version 2. When you configure the static RP, the RP address that you select for a particular group must be consistent across all routers in a multicast domain.

A static configuration is simple and convenient. However, if the statically defined RP router becomes unreachable, there is no automatic failover to another RP router. To remedy this problem, you can use anycast RP.

Anycast RP

Anycast means that multiple RP routers share the same unicast IP address. Anycast addresses are advertised by the routing protocols. Packets sent to the anycast address are sent to the nearest RP with this address. Anycast addressing is a generic concept and is used in PIM sparse mode to add load balancing and service reliability to RPs.

Having a single active RP per multicast group is much the same as having a single server providing any service. All traffic converges on this single point, although other servers are sitting idle, and convergence is slow when the resource fails. In multicast specifically, there might be closer RPs on the shared tree, so the use of a single RP is suboptimal.

When anycast RP is configured, the shared address is used in the RP-to-group mapping. This allows multicast groups to have multiple active RPs in a PIM domain. However, the RPs must use some protocol to synchronize the active source information so that the active RP for each group is known to all RPs.

There are two methods for RP active source synchronization in anycast RP, one using the Multicast Source Discovery Protocol (MSDP) and the other using PIM itself.

When MSDP is used with PIM sparse mode, anycast RP provides a faster failover rate than auto-RP or a bootstrap router. However, MSDP only works for IPv4. When PIM alone is used for anycast RP, the solution works for both IPv4 and IPv6.

For more information about configuring static RPs, see “Configuring Static RPs” on page 227. For more information about configuring anycast RP, see “Configuring Auto-RP” on page 231 and “Example: Configuring Anycast RP” on page 248.

Auto-RP

You can configure a more dynamic way of assigning RPs in a multicast network by means of auto-RP. When you configure auto-RP for a router, the router learns the address of the RP in the network automatically and has the added advantage of operating in PIM version 1 and version 2.

Although auto-RP is a nonstandard (non-RFC-based) function that typically uses dense mode PIM to advertise control traffic, it provides an important failover advantage that simple static RP assignment does not. You can configure multiple routers as RP candidates. If the elected RP stops operating, one of the other preconfigured routers takes over the RP functions. This capability is controlled by the auto-RP mapping agent.

For more information, see “Configuring Auto-RP” on page 231.

Bootstrap Router

To determine which router is the RP, all routers within a PIM sparse-mode domain collect bootstrap messages. A PIM sparse-mode domain is a group of routers that all share the same RP router. The domain's bootstrap router originates bootstrap messages, which are sent hop by hop within the domain. The routers use bootstrap messages to distribute RP information dynamically and to elect a bootstrap router when necessary.

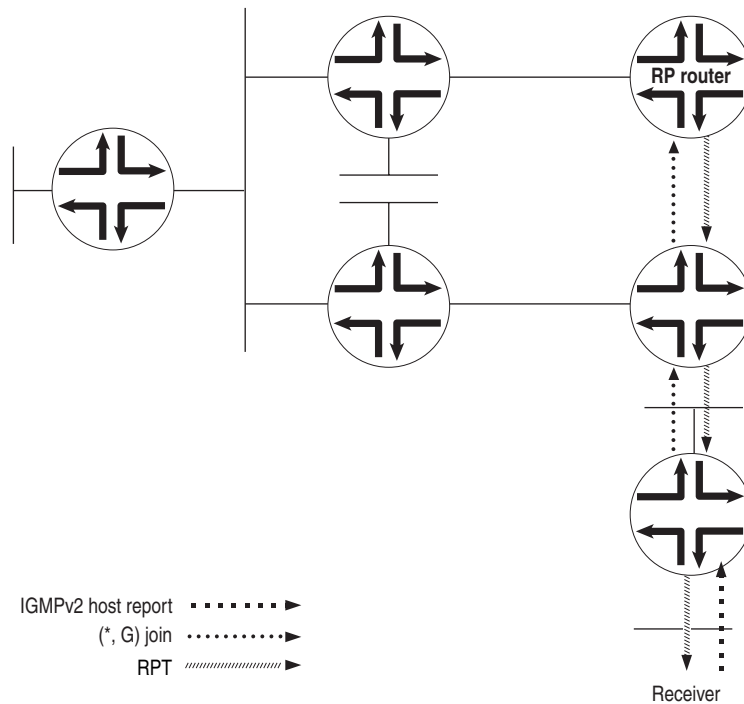
For more information, see “Configuring Bootstrap Properties” on page 228.

Building an RPT Between RP and Receivers

The RPT is the path between the RP and receivers (hosts) in a multicast group (see Figure 15). The RPT is built by means of a PIM join message from a receiver's DR:

1. A receiver sends a request to join group (G) in an Internet Group Management Protocol (IGMP) host membership report. A PIM sparse-mode router, the receiver's DR, receives the report on a directly attached subnet and creates an RPT branch for the multicast group of interest.
2. The receiver's DR sends a PIM join message to its RPF neighbor, the next-hop address in the RPF table, or the unicast routing table.
3. The PIM join message travels up the tree, and is multicast to the ALL-PIM-ROUTERS group (224.0.0.13). Each router in the tree finds its RPF neighbor by using either the RPF table or the unicast routing table. This is done until the message reaches the RP and forms the RPT. Routers along the path set up the multicast forwarding state to forward requested multicast traffic back down the RPT to the receiver.

Figure 15: Building an RPT Between RP and Receiver



g017113

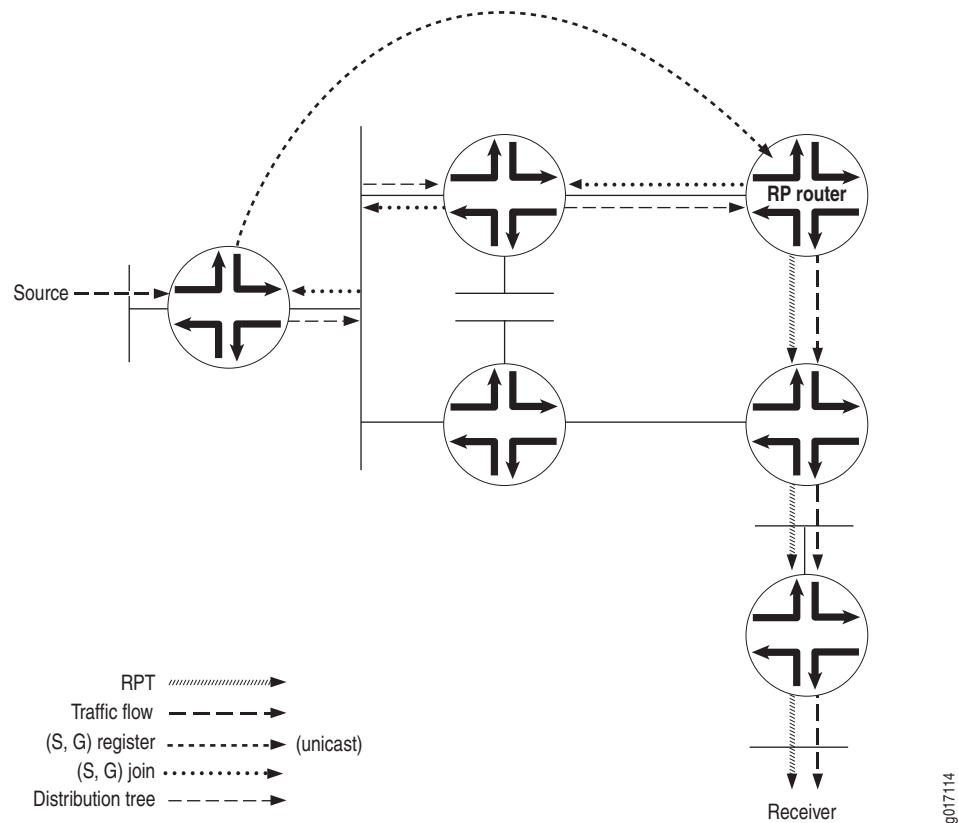
PIM Sparse-Mode Source Registration

The RPT is a unidirectional tree, permitting traffic to flow down from the RP to the receiver in one direction. For multicast traffic to reach the receiver from the source, another branch of the distribution tree called the SPT needs to be built from the source's DR to the RP.

The SPT is created in the following way:

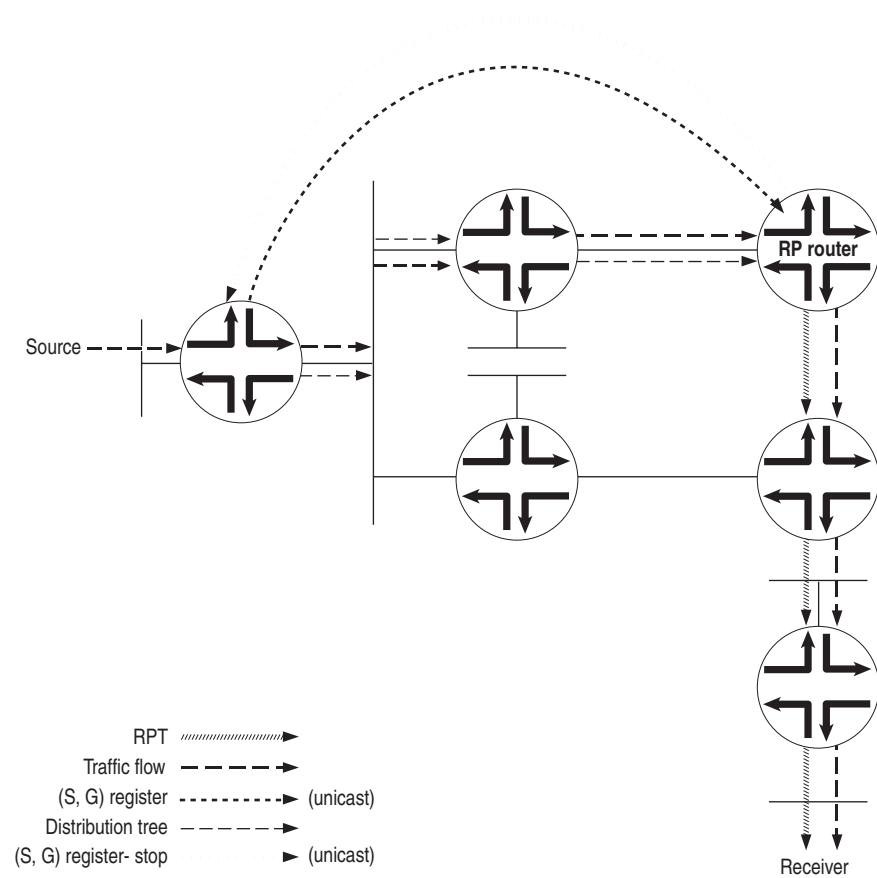
1. The source becomes active, sending out multicast packets on the LAN to which it is attached. The source's DR receives the packets and encapsulates them in a PIM register message, which it sends out to the RP router (see Figure 16).
2. When the RP router receives the PIM register message from the source, it sends a PIM join message back to the source.

Figure 16: PIM Register Message and PIM Join Message Exchanged



3. The source's DR receives the PIM join message and begins sending traffic down the SPT toward the RP router (see Figure 17).
4. Once traffic is received by the RP router, it sends a register stop message to the source's DR to stop the register process.

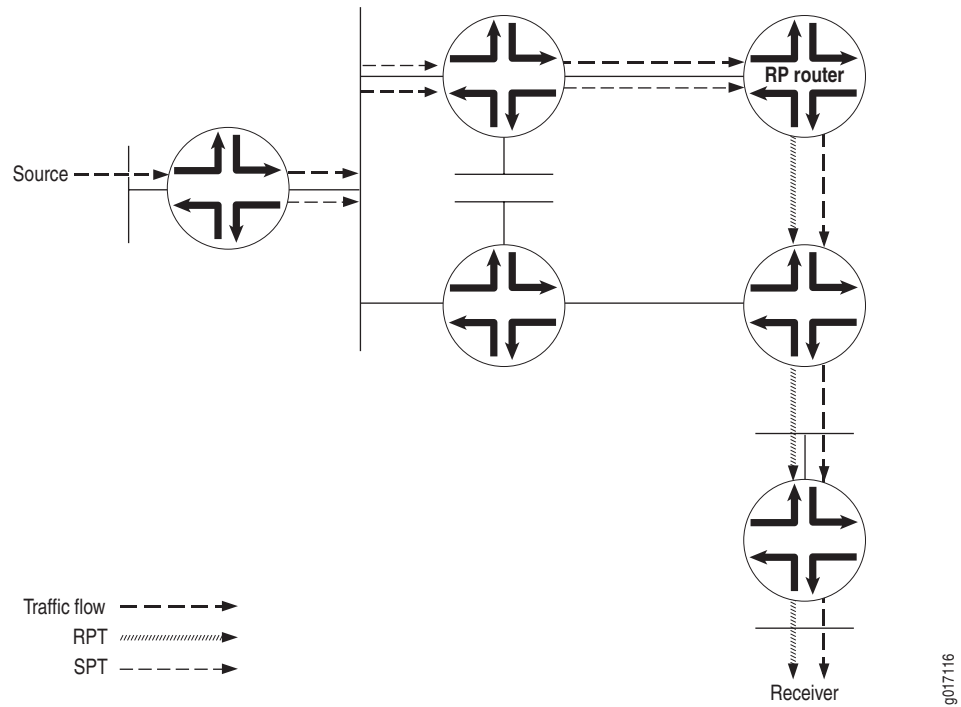
Figure 17: Traffic Sent from the Source to the RP Router



9017115

- The RP router sends the multicast traffic down the RPT toward the receiver (see Figure 18).

Figure 18: Traffic Sent from the RP Router Toward the Receiver



PIM Sparse-Mode SPT Cutover

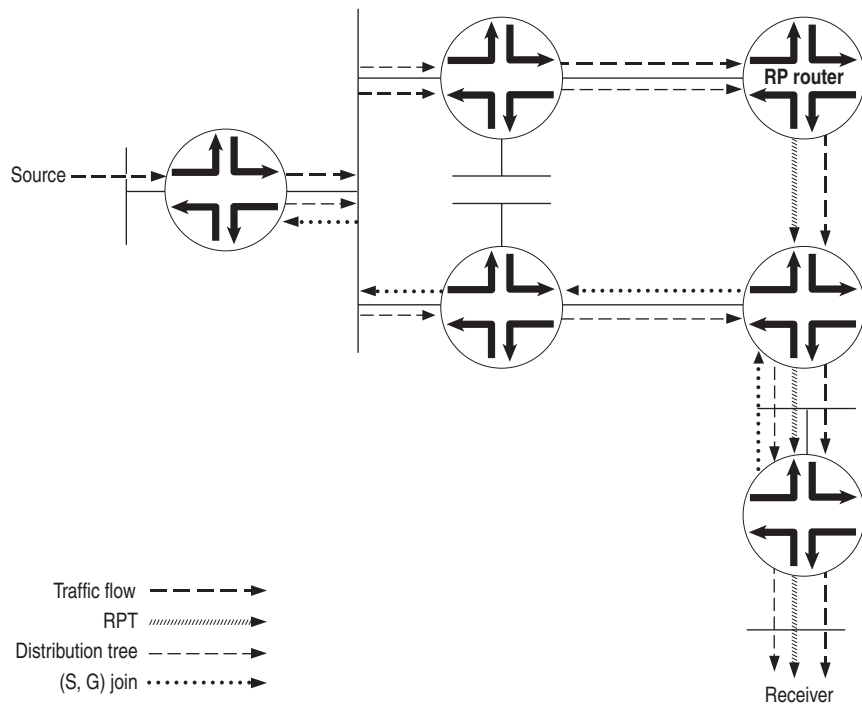
The RPT is not always the most direct path for delivering multicast traffic to a receiver. In many cases, a direct SPT from the last-hop router to the source is a better way to receive a multicast stream.

SPT Cutover

Instead of continuing to use the SPT to the RP and the RPT toward the receiver, a direct SPT is created between the source and the receiver in the following way:

1. Once the receiver's DR receives the first multicast packet from the source, the DR sends a PIM join message to its RPF neighbor (see Figure 19).
2. The source's DR receives the PIM join message, and an additional (S,G) state is created to form the SPT.
3. Multicast packets from that particular source begin coming from the source's DR and flowing down the new SPT to the receiver's DR. The receiver's DR is now receiving two copies of each multicast packet sent by the source—one from the RPT and one from the new SPT.

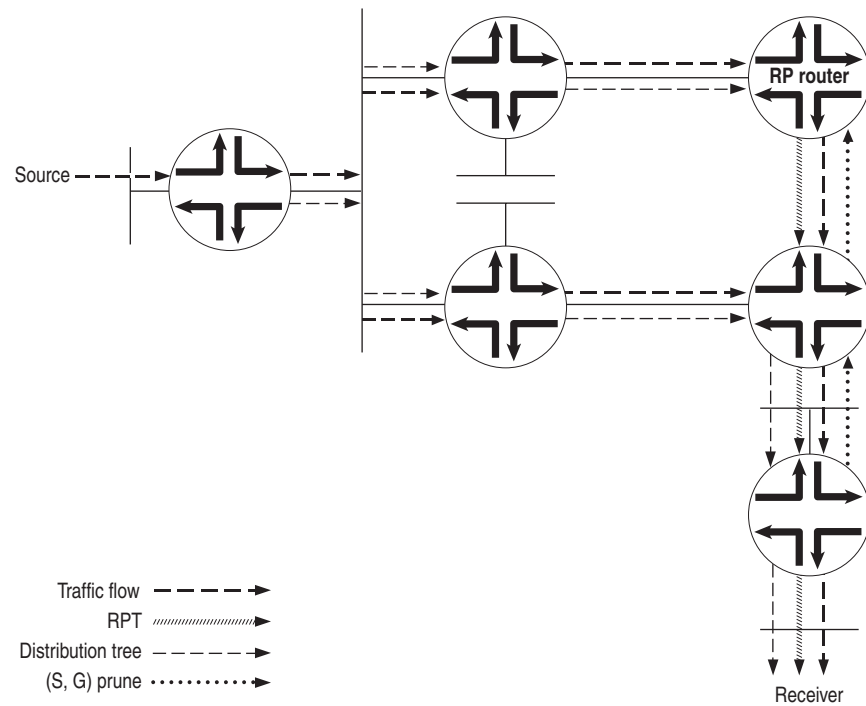
Figure 19: Receiver DR Sends a PIM Join Message to the Source



g017117

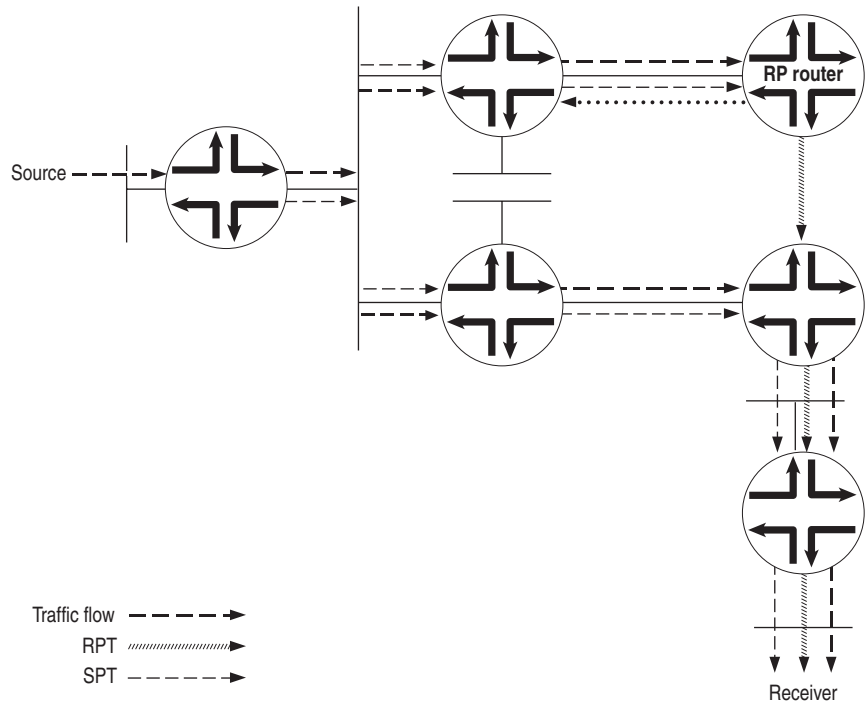
- To stop duplicate multicast packets, the receiver's DR sends a PIM prune message toward the RP router, letting it know that the multicast packets from this particular source coming in from the RPT are no longer needed (see Figure 20).

Figure 20: PIM Prune Message is Sent from the Receiver's DR Toward the RP Router



- The PIM prune message is received by the RP router, and it stops sending multicast packets down to the receiver's DR. The receiver's DR is getting multicast packets only for this particular source over the new SPT. However, multicast packets from the source are still arriving from the source's DR toward the RP router (see Figure 21).

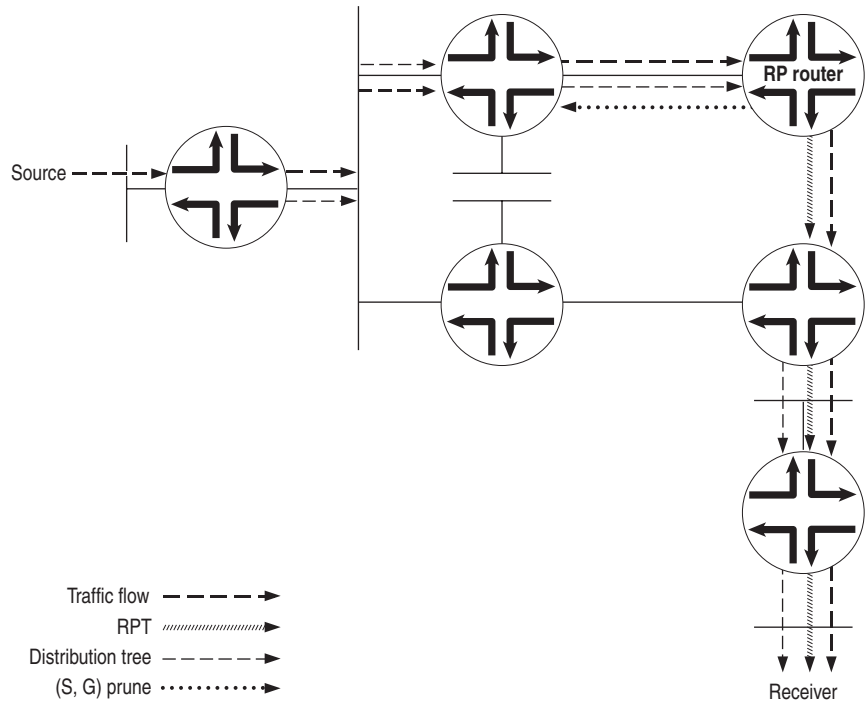
Figure 21: RP Router Receives PIM Prune Message



g017119

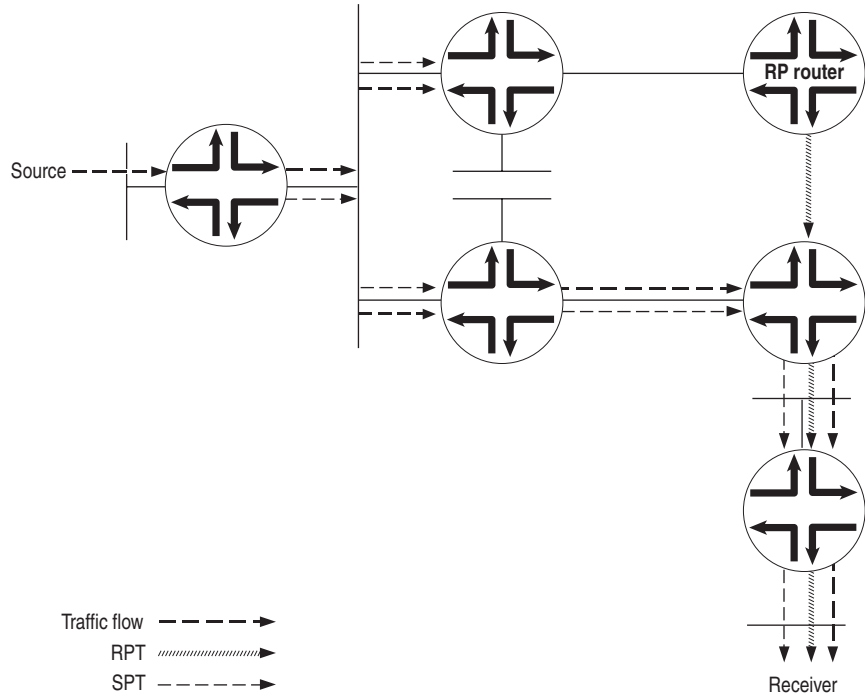
- To stop the unneeded multicast packets from this particular source, the RP router sends a PIM prune message to the source's DR (see Figure 22).

Figure 22: RP Router Sends a PIM Prune Message to the Source DR



- The receiver's DR now receives multicast packets only for the particular source from the SPT (see Figure 23).

Figure 23: Source's DR Stops Sending Duplicate Multicast Packets Toward the RP Router



g017121

SPT Cutover Control

In some cases, the last-hop router should stay on the shared tree to the RP and *not* transition to a direct SPT to the source. You might not want the last-hop router to transition when, for example, a low-bandwidth multicast stream is forwarded from the RP to a last-hop router. All routers between last hop and source must maintain and refresh the SPT state. This can become a resource-intensive activity that does not add much to the network efficiency for a particular pair of source and multicast group addresses.

In these cases, you configure an SPT threshold policy on the last-hop router to control the transition to a direct SPT. An SPT cutover threshold of infinity applied to a source-group address pair means the last-hop router will *never* transition to a direct SPT. For all other source-group address pairs, the last-hop router transitions immediately to a direct SPT rooted at the source DR.

PIM SSM

PIM source-specific multicast (SSM) uses a subset of PIM sparse mode and IGMP version 3 (IGMPv3) to allow a client to receive multicast traffic directly from the source. PIM SSM uses the PIM sparse-mode functionality to create an SPT between the receiver and the source, but builds the SPT without the help of an RP.

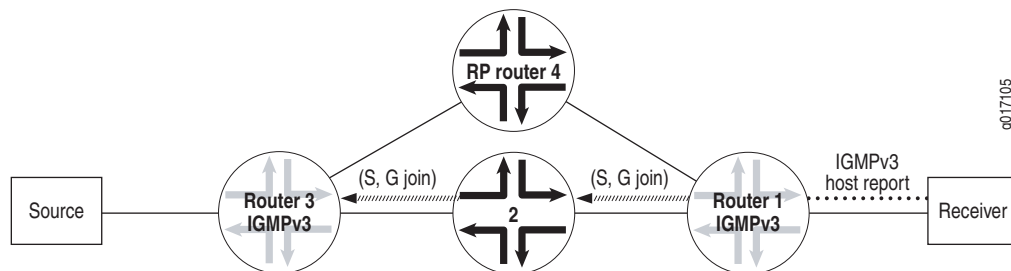
By default, the SSM group multicast address is limited to the IP address range 232.0.0.0 to 232.255.255.255. However, you can extend SSM operations into another Class D range by including the `address` statement at the `[edit routing-options multicast ssm-groups]` hierarchy level.

An SSM-configured network has distinct advantages over a traditionally configured PIM sparse-mode network. There is no need for shared trees or RP mapping (no RP is required), or for RP-to-RP source discovery through MSDP.

Deploying SSM is easy. You need only configure PIM sparse mode on all router interfaces and issue the necessary SSM commands, including specifying IGMPv3 on the receiver's LAN. If PIM sparse mode is not explicitly configured on both the source and group member interfaces, multicast packets are not forwarded. Source lists, supported in IGMPv3, are used in PIM SSM. As sources become active and start sending multicast packets, interested receivers in the SSM group receive the multicast packets.

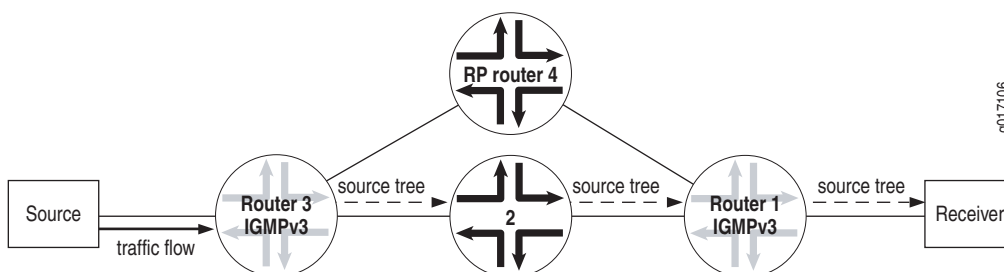
In a PIM SSM-configured network, a host subscribes to an SSM channel (by means of IGMPv3), announcing a desire to join group G and source S (see Figure 24). The directly connected PIM sparse-mode router, the receiver's DR, sends an (S,G) join message to its RPF neighbor for the source. Notice in Figure 24 that the RP is not contacted in this process by the receiver, as would be the case in normal PIM sparse-mode operations.

Figure 24: Receiver Announces Desire to Join Group G and Source S



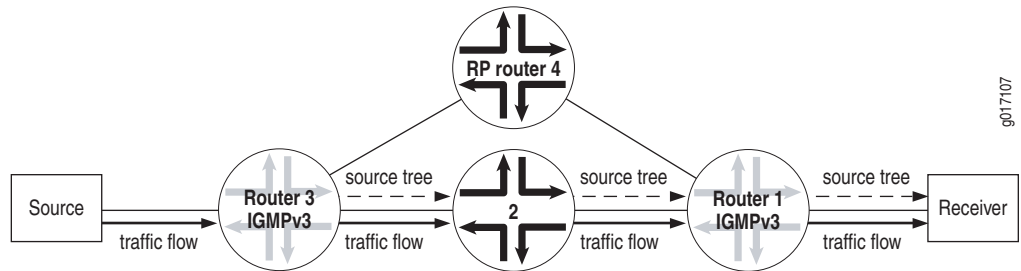
The (S,G) join message initiates the source tree, then builds it out hop by hop until it reaches the source. In Figure 25, the source tree is built across the network to Router 3, the last-hop router connected to the source.

Figure 25: Router 3 (Last-Hop Router) Joins the Source Tree



Using the source tree, multicast traffic is delivered to the subscribing host (see Figure 26).

Figure 26: The (S,G) State Is Built Between the Source and the Receiver



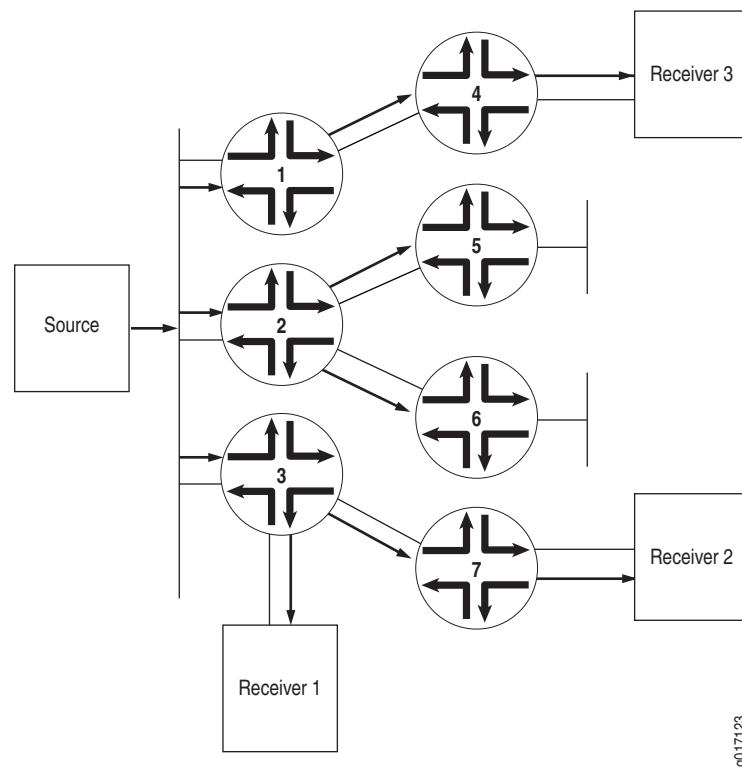
To configure additional SSM groups, include the `ssm-groups` statement at the [edit routing-options multicast] hierarchy level.

For more information about PIM SSM, see “Example: Configuring PIM SSM on a Network” on page 146.

PIM Dense Mode

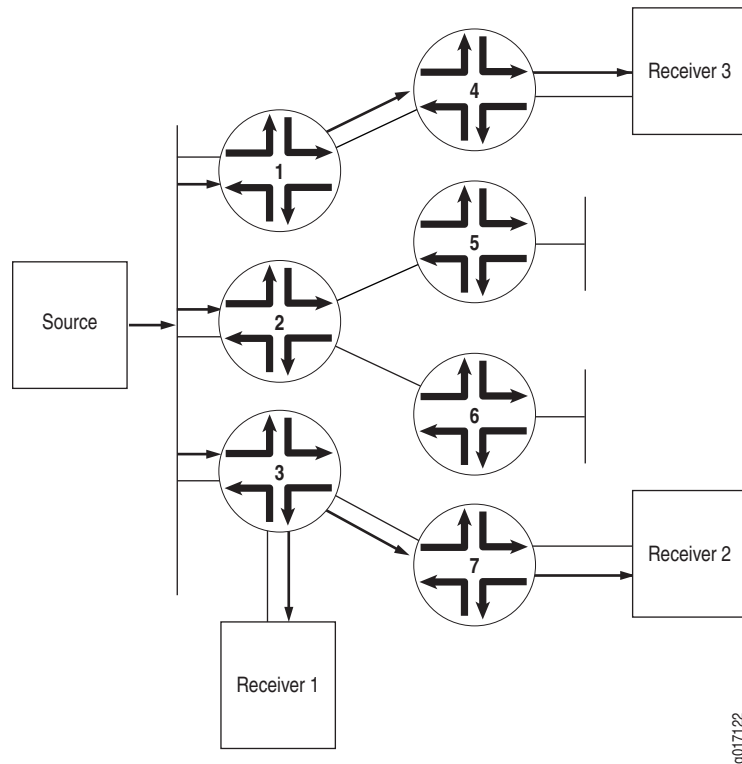
Unlike sparse mode, in which data is forwarded only to routers sending an explicit request, dense mode implements a *flood-and-prune* mechanism, similar to DVMRP. In PIM dense mode, there is no RP. A router receives the multicast data on the interface closest to the source, then forwards the traffic to all other interfaces (see Figure 27).

Figure 27: Multicast Traffic Flooded from the Source Using PIM Dense Mode



Flooding occurs periodically. It is used to refresh state information, such as the source IP address and multicast group pair. If the router has no interested receivers for the data, and the OIL becomes empty, the router sends a prune message upstream to stop delivery of multicast traffic (see Figure 28).

Figure 28: Prune Messages Sent Back to the Source to Stop Unwanted Multicast Traffic



PIM Sparse-Dense Mode

Sparse-dense mode, as the name implies, allows the interface to operate on a per-group basis in either sparse or dense mode. A group specified as dense is not mapped to an RP. Instead, data packets destined for that group are forwarded by means of PIM dense-mode rules. A group specified as sparse is mapped to an RP, and data packets are forwarded by means of PIM sparse-mode rules.

For information about PIM sparse-mode and PIM dense-mode rules, see “PIM Sparse Mode” on page 186 and “PIM Dense Mode” on page 203.

RP Mapping with Anycast RP

For the purposes of load balancing and redundancy, you can configure anycast RP. When you configure anycast RP, you bypass the restriction of having one active RP per multicast group, and instead deploy multiple RPs for the same group range. The RP routers share one unicast IP address. Sources from one RP are known to other RPs that use MSDP. Sources and receivers use the closest RP, as determined by the interior gateway protocol (IGP).

You can use anycast RP within a domain to provide redundancy and RP load sharing. When an RP goes down, sources and receivers are taken to a new RP by means of unicast routing.

Anycast RP is defined in Internet draft `draft-ietf-mboned-anycast-rp-08.txt`, *Anycast RP Mechanism Using PIM and MSDP*. To access Internet RFCs and drafts, go to the IETF Web site at <http://www.ietf.org>.

We recommend a static RP mapping with anycast RP and a bootstrap router (BSR) with auto-RP configuration because static mapping provides all the benefits of a bootstrap router and auto-RP without the complexity of the full BSR and auto-RP mechanisms.

See also “Example: Configuring Anycast RP” on page 248.

Multicast over Layer 3 VPNs

In the unicast environment for Layer 3 virtual private networks (VPNs), all VPN state information is contained within the provider edge (PE) routers. However, with multicast for Layer 3 VPNs, PIM adjacencies are established in one of the following ways:

- You can set PIM adjacencies between the customer edge (CE) router and the PE router through a VPN routing and forwarding (VRF) instance at the `[edit routing-instances instance-name protocols pim]` hierarchy level. You must include the new `vpn-group-address` statement at this hierarchy level, specifying a multicast group. The RP listed in the VRF-instance is the VPN customer RP (C-RP).
- You can also set the master PIM instance and the PE’s IGP neighbors by configuring statements at the `[edit protocols pim]` hierarchy level. You must add the multicast group specified in the VRF instance to the master PIM instance. The set of master PIM adjacencies throughout the service provider network makes up the forwarding path that becomes an RP tree rooted at the service provider RP (SP-RP). Therefore, provider (P) routers within the provider core must maintain multicast state information for the VPNs.

For this configuration to work properly, you need two types of RP routers for each VPN:

- A VPN C-RP—An RP router located somewhere within the customer VPN.
- An SP-RP—An RP router located within the service provider network.



NOTE: A PE router can act as an SP-RP or the VPN C-RP of a VPN. However, when auto-RP and BSR are used, the PE cannot be a C-RP. It can, however, learn another router as C-RP by means of the auto-RP or BSR protocols.

For more information about configuring multicast for Layer 3 VPNs, see “Configuring Multicast for Layer 3 VPNs” on page 241. For multicast Layer 3 VPN examples, see “Example: Configuring PIM Sparse Mode over Layer 3 VPNs” on page 257.

Tunnel Services PICs and Multicast

On Juniper Networks routers, data packets are encapsulated and de-encapsulated into tunnels by means of hardware and not the software running on the router’s processor. The hardware used to create tunnel interfaces is a Tunnel Services Physical Interface Card (PIC). All RP routers and IP version 4 (IPv4) PIM sparse-mode DRs connected to a source require a Tunnel Services PIC.

In PIM sparse mode, the source DR takes the initial multicast packets and encapsulates them in PIM register messages. It then unicasts them to the PIM sparse-mode RP router, where the PIM register message is de-encapsulated.

When a router is configured as a PIM sparse-mode RP router (by specifying an address using the `address` statement at the [edit protocols pim rp local] hierarchy level) and a Tunnel PIC is present on the router, a PIM register de-encapsulation interface, or `pd` interface, is automatically created. The `pd` interface receives PIM register messages and de-encapsulates them by means of the hardware.

If PIM sparse mode is enabled on any router (potentially a PIM sparse-mode source DR) and a Tunnel Services PIC is present on the router, a PIM register encapsulation interface, or `pe` interface, is automatically created for each RP address that is used to encapsulate source data packets and send them to respective RP addresses on the PIM DR as well as the PIM RP. The `pe` interface receives PIM register messages and encapsulates them by means of the hardware.



NOTE: Do not confuse the configurable `pe` and `pd` hardware interfaces with the nonconfigurable `pime` and `pimd` software interfaces. Both pairs encapsulate and de-encapsulate multicast packets, and are created automatically; however, the `pe` and `pd` interfaces only appear if a Tunnel Services PIC is present. The `pime` and `pimd` interfaces are not useful in situations requiring the `pe` and `pd` interfaces.

If the source DR is the RP, then there is no need for PIM register messages and consequently no need for a Tunnel Services PIC to be present.

When PIM sparse mode is used with IP version 6 (IPv6), a Tunnel PIC is required on the RP, but not on the IPv6 PIM DR. The lack of a Tunnel PIC requirement on the IPv6 DR applies only to IPv6 PIM sparse mode and should not be confused with IPv4 PIM sparse-mode requirements.

Table 7 shows the complete matrix of IPv4 and IPv6 PIM Tunnel PIC requirements.

Table 7: Tunnel PIC Requirements for IPv4 and IPv6 Multicast

IP Version:	Tunnel PIC on RP	Tunnel PIC on DR
IPv4	Yes	Yes
IPv6	Yes	No

Filtering Multicast Messages

Multicast sources and routers generate a considerable number of control messages, especially when using PIM sparse mode. These messages form distribution trees, locate RPs and DRs, and transition from one type of tree to another. In most cases, this multicast messaging system operates transparently and efficiently. However, in some configurations, more control over the sending and receiving of multicast control messages is necessary.

You can configure two types of multicast filtering to control the sending and receiving of multicast control messages.

This section discusses three types of multicast filtering. The last two require configuration:

- Filtering MAC Addresses on page 207
- Filtering RP/DR Register Messages on page 208
- Filtering MSDP SA Messages on page 209

Filtering MAC Addresses

When a router is exclusively configured with multicast protocols on an interface, multicast sets the interface media access control (MAC) filter to multicast promiscuous mode, and the number of multicast groups is unlimited. However, when the router is not exclusively used for multicasting and other protocols such as Open Shortest Path First (OSPF), Routing Information Protocol version 2 (RIPv2), or Network Time Protocol (NTP) are configured on an interface, each of these protocols individually requests the interface to program the MAC filter to pick up their respective multicast group alone. In this case, without multicast configured on the interface, the maximum number of multicast MAC filters is limited to 20. For example, the maximum number of interface MAC filters for protocols such as OSPF (multicast group 224.0.0.5) is 20, unless a multicast protocol is also configured on the interface.

No configuration is necessary for MAC filters.

Filtering RP/DR Register Messages

You can filter PIM register messages sent from the DR or to the RP. The PIM RP keeps track of all active sources in a single PIM sparse mode domain. In some cases, more control over which sources an RP knows about, or which sources a DR tells other RPs about, is desired. A high degree of control over PIM register messages is provided by RP/DR register message filtering. Message filtering also prevents unauthorized groups and sources from registering with an RP router.

Register messages that are filtered at a DR are not sent to the RP, but the sources are available to local users. Register messages that are filtered at an RP arrive from source DRs, but are ignored by the router. Sources on multicast group traffic can be limited or directed by using RP or DR register message filtering alone or in combination.

If the action of the register filter policy is to discard the register message, the router should send a register-stop message to the DR. These register-stop messages are throttled to prevent malicious users from triggering them on purpose to disrupt the routing process.

Multicast group and source information is encapsulated inside unicast IP packets. This feature allows the router to inspect the multicast group and source information before sending or accepting the PIM register message.

Incoming register messages to an RP are passed through the configured register message filtering policy before any further processing. If the register message is rejected, the RP router sends a register stop message to the DR. When the DR receives the register stop message, the DR stops sending register messages for the filtered groups and sources to the RP. Two fields are used for register message filtering:

- Group multicast address
- Source address

The syntax of the existing policy statements are used to configure the filtering on these two fields. The **router-filter** statement is useful for multicast group address filtering and the **source-address-filter** statement is useful for source address filtering. In most cases, the action will be to **reject** the register messages, but more complex filtering policies are possible.

Filtering cannot be performed on other header fields, such as DR address, protocol, or port. In some configurations, an RP might not send register-stop messages when the policy action is to discard the register messages. This has no effect on the operation of the feature, but the router will continue to receive register messages.

When anycast RP is configured, register messages can be sent or received by the RP. All the RPs in the anycast RP set should have the same RP register message filtering policies configured; otherwise, it might be possible to circumvent the filtering policy. For more information on anycast RP, see “RP Mapping with Anycast RP” on page 205.

For more information on filtering RP/DR register messages, see “Configuring RP/DR Register Message Filtering” on page 236 and “Example: Configuring RP/DR Register Message Filters” on page 253.

Filtering MSDP SA Messages

Along with applying MSDP source active (SA) filters on all external MSDP sessions (in and out) to prevent SAs for groups and sources from leaking in and out of the network, you need to apply BSR filters. Applying a BSR filter to the boundary of a network prevents foreign BSR messages (which announce RP addresses) from leaking into your network. Since the routers in a PIM sparse-mode domain should know the address of only one RP router, having more than one in the network can create problems. See “Example: Configuring PIM BSR Filters” on page 252 for a sample filter configuration.

If you did not use multicast scoping to create boundary filters for all customer-facing interfaces, you might want to use PIM join filters. Multicast scopes prevent the actual multicast data packets from flowing in or out of an interface. PIM join filters prevent PIM sparse-mode state from being created in the first place. Since PIM join filters apply only to the PIM sparse-mode state, it might be more beneficial to use multicast scoping to filter the actual data.

For more information, see “Multicast Scoping Overview” on page 133 and “Example: Configuring PIM Join Filters” on page 252.



NOTE: When you apply firewall filters, firewall action modifiers, such as log, sample, and count, work only when you apply the filter on an inbound interface. The modifiers do not work on an outbound interface.

Embedded RP for IPv6 Multicast

Global IPv6 multicast between routing domains has been possible only with SSM because there is no way to convey information about IPv6 multicast RPs between PIM sparse mode RPs. In IPv4 multicast networks this information is conveyed between PIM RPs using MSDP, but there is no IPv6 support in current MSDP standards. IPv6 uses the concept of an embedded RP to resolve this issue without requiring SSM. This feature embeds the RP address in an IPv6 multicast address.

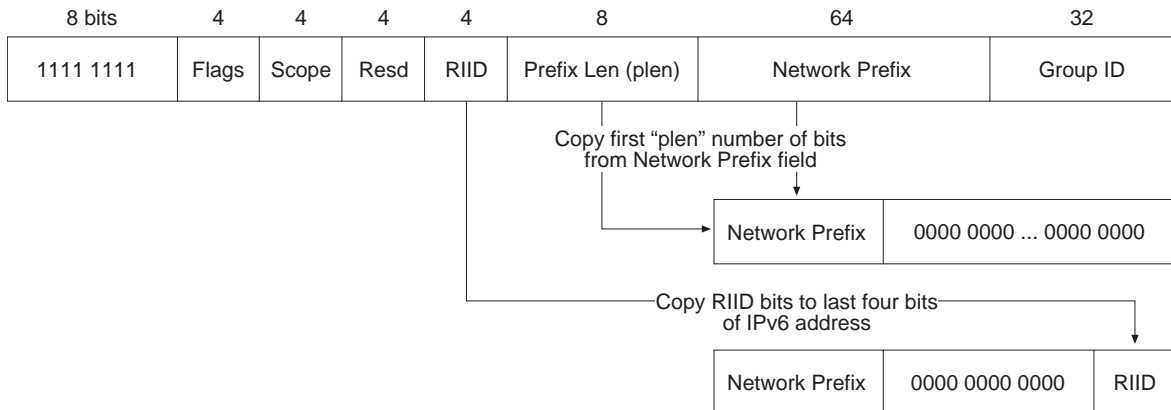
All IPv6 multicast addresses begin with 8 1-bits (**1111 1111**) followed by a 4-bit flag field normally set to **0011**. The flag field is set to **0111** when embedded RP is used. Then the low-order bits of the normally reserved field in the IPv6 multicast address carry the 4-bit RP interface identifier (RIID).

When the IPv6 address of the RP is embedded in a unicast-prefix-based any-source multicast (ASM) address, all of the following conditions must be true:

- The address must be an IPv6 multicast address and have **0111** in the flags field (that is, the address is part of the prefix **FF70::/12**).
- The 8-bit prefix length (plen) field must not be all 0. An all 0 plen field implies that SSM is in use.
- The 8-bit prefix length field value must not be greater than 64, which is the length of the network prefix field in unicast-prefix-based ASM addresses.

The routing platform derives the value of the interdomain RP by copying the prefix length field number of bits from the 64-bit network prefix field in the received IPv6 multicast address to an empty 128-bit IPv6 address structure and copying the last bits from the 4-bit RIID. For example, if the prefix length field bits have the value 32, then the routing platform copies the first 32 bits of the IPv6 multicast address network prefix field to an all-0 IPv6 address and appends the last four bits determined by the RIID. See Figure 29 for an illustration of this procedure.

Figure 29: Extracting the Embedded RP IPv6 Address



For example, the administrator of IPv6 network 2001:DB8::/32 sets up an RP for the 2001:DB8:BEEF:FEED::/96 subnet. In that case, the received embedded RP IPv6 ASM address has the form:

FF70:y40:2001:DB8:BEEF:FEED::/96

and the derived RP IPv6 address has the form:

2001:DB8:BEEF:FEED::y

where y is the RIID (y cannot be 0).

When configured, the routing platform checks for embedded RP information in every PIM join request received for IPv6. The use of embedded RP does not change the processing of IPv6 multicast and RPs in any way, except that the embedded RP address is used if available and selected for use. There is no need to specify the IPv6 address family for embedded RP configuration because the information can be used only if IPv6 multicast is properly configured on the routing platform.

The following receive events trigger extraction of an IPv6 embedded RP address on the routing platform:

- Multicast Listener Discovery (MLD) report for an embedded RP multicast group address
- PIM join message with an embedded RP multicast group address
- Static embedded RP multicast group address associated with an interface
- Packets sent to an embedded RP multicast group address received on the DR

The embedded RP node discovered through these events is added if it does not already exist on the routing platform. The routing platform chooses the embedded RP as the RP for a multicast group before choosing an RP learned through BSR or a statically configured RP. The embedded RP is removed whenever all PIM join states using this RP are removed or the configuration changes to remove the embedded RP feature.

